Class Size Reduction, Teacher Quality, and Academic Achievement in California Public Elementary Schools

• • •

Christopher Jepsen
Steven Rivkin

2002

PUBLIC POLICY INSTITUTE OF CALIFORNIA

Library of Congress Cataloging-in-Publication Data
Jepsen, Christopher.
Class size reduction, teacher quality, and academic achievement in California public elementary schools / Christopher Jepsen, Steven Rivkin.
p. cm.
Includes bibliographical references.
ISBN: 1-58213-058-2
1. Class size—California. 2. Elementary school teachers—Training of—California. 3. Academic achievement—California. I. Rivkin, Steven G. (Steven Gary), 1961– II. Title.

LB3013.2 .J47 2002
371.2'51—dc21 2002069687

Foreword

Intuitively, class size reduction is a good idea. Parents support it because it means that their children will receive more individual attention from teachers. Teachers like it for the same reason and also because it creates a more manageable workload. It is generally assumed that the fewer students in a class, the better they will learn and the higher they will score on nationwide tests—currently one of the most common measures of student achievement.

Nevertheless, several things could go wrong with mandated class size reduction (CSR). This report presents evidence that the potential success or failure of such a reform may depend largely on how it is implemented and how teachers and administrators respond. If it is implemented quickly on a large scale, such a program may run into serious problems. In the following pages, Christopher Jepsen and Steven Rivkin analyze some of the things that went right and some that did not after California passed its CSR law in 1996.

One problem was the sweeping nature of the legislation, which called for a one-third reduction in class size for kindergarten through third grade in every public elementary school in the state. This resulted in an immediate need for thousands of new teachers. But where were they to come from? And what qualifications would they have? In 1997, the first year after the schools had scrambled to hire new teachers, nearly one-quarter of the teaching workforce in California had one year of experience or less.

The authors also point to a more serious problem. Many teachers in economically disadvantaged communities left their schools to fill vacancies created in other schools. Thus, many schools had to fill not only the positions created by their own efforts to reduce class size but also positions vacated by departing teachers.

The analysis reveals that the effects of class size reduction extended beyond the third grade, leading to lower achievement in the fifth grade among schools with a high percentage of black students. Evidently, many fourth- and fifth-grade teachers chose to move to the earlier grades with fewer students, leaving many schools with the need to hire new teachers in other grades as well as K–3. In fact, the analysis indicates that schools that had trouble hiring experienced teachers before CSR are the very schools that did not appear to benefit from CSR.

On the positive side, the authors find that, all else equal, smaller classes raise student achievement. Reducing class size by 10 students raises the percentage of third-grade students who exceed national median test scores by about 4 percentage points in mathematics and 3 percentage points in reading. However, all else is not equal when you hire thousands of new teachers. Separating out the effects of new teachers from the effects of smaller classes, Jepsen and Rivkin find that having a new teacher reduces the percentage of students who exceed the national median by roughly 3 percentage points in both mathematics and reading. So in many schools, class size reduction has meant zero gain, at least in the short run.

The bottom line is that California has implemented a $1.6-billion-a-year program that is apparently yielding only modest gains. It also appears that some students and schools have actually been hurt rather than helped by class size reduction.

And now the cost of the reform is becoming a problem itself. California implemented CSR during prosperous times, when it had a strong budget surplus. The state also picked up most of the cost. Today, California is suffering from a $20 billion shortfall, and school districts find themselves shouldering more of the fiscal burden. Faced with difficulty in meeting essential costs, such as teacher salaries, maintenance, and materials for the classroom, some districts are choosing to leave the program, preferring larger classes to cutbacks in other areas.

CSR does not appear to have been the silver bullet many expected it to be. If it remains in place, one obvious policy response at this point would be to help beginning teachers adapt to the classroom more quickly and effectively and gain the essential training they need. An ongoing concern that also needs to be addressed is the large disparity across schools in teacher qualifications and student achievement. There is sufficient evidence to suggest that improving the quality of the teaching staff in any school will go a long way toward improving student achievement.

According to the authors, another lesson that other states might learn from the California experience is this: "A better approach to class size reduction would have been to reduce class sizes in a subset of schools each year, starting with low-performing schools serving high-poverty populations. This would have limited the departure of teachers for newly created jobs in suburban schools, lessened the overall competition for new teachers, and reduced inequality in academic performance."

No single reform will improve California's K–12 system. As we have learned from the class size reduction program, unintended consequences can offset gains as well as create whole new problems. Much can be learned, however, from incremental policy changes for a limited number of school districts before scaling up to statewide implementation. There is much still to be done, and the lessons from CSR will make the next steps more informed.

David W. Lyon
President and CEO
Public Policy Institute of California

Summary

In response to widespread dissatisfaction with the public schools, California has implemented a number of educational reforms over the past decade. Perhaps the most dramatic, and certainly the most costly, was the passage of the class size reduction (CSR) law in the summer of 1996. This legislation aimed to reduce average class size in grades kindergarten through third grade by roughly one-third, from 30 students to 20, at an annual cost of over $1 billion.

Educators and policymakers were hopeful that the gains attributed to smaller class sizes in an experiment conducted in Tennessee would translate into large gains in achievement following California's statewide class size reduction. However, the hiring of large numbers of inexperienced teachers in response to CSR had the potential to offset the direct benefits of smaller classes, particularly for schools in economically disadvantaged communities that had extreme staffing difficulties before class size reduction.

Because test scores before CSR are not available for most schools, the total effect of CSR will never be known. However, this report attempts to provide some limited but very important answers to the following questions:

1. What were the effects of CSR on teacher experience, certification, and education? Were some schools affected more than others?
2. How did CSR affect student achievement in third grade? What were the benefits of smaller classes? What were the effects of new teachers?
3. Are the benefits of smaller classes concentrated among a subset of students, or did all schools benefit equally from CSR? For example, many researchers find that low-income and nonwhite students benefit more from smaller classes than do other students.

Student Diversity

Immigration and other factors led to substantial demographic changes in California public elementary schools during the 1990s. No single ethnic group constitutes a majority of students: However, Hispanics are the most populous group. At the same time, large percentages of students are attending schools where more than half the student population is enrolled in free or reduced-price (i.e., subsidized) lunch programs. There are substantial differences across the state in the distribution of students by race/ethnicity and by the percentage enrolled in subsidized lunch programs. For example, urban schools have higher percentages of nonwhite students and low-income students than do suburban or rural schools. Such differences in demographic composition, even among the largest districts in the state, raise the possibility that CSR may have had dramatically different effects across the state, depending largely on the strength of the relationship between teacher characteristics and student demographic composition.

Teacher Characteristics

Numerous reports have documented the change in teacher characteristics following the implementation of CSR. Although our description of California elementary school teachers covers much of the same ground as earlier work, there are several key additions. First, rather than combining all teachers with less than three years of experience into a single category, we create separate categories for first- and second-year teachers and group all remaining teachers with at least two years of experience into a single category. Second, earlier studies examined differences in teacher characteristics by community type, student racial/ethnic composition, or student income, but they did not consider the distribution of teacher characteristics for combinations of these factors. As the distribution of students within each subsidized lunch category (0 to 25 percent, 25 to 50 percent, 50 to 75 percent, 75 to 100 percent) differs by race, the distribution of teacher characteristics likely differs as well. However, these teacher differences would be overlooked if one were focusing solely on either race/ethnicity or income categories.

CSR led to a dramatic increase in the percentages of inexperienced and uncertified teachers. In 1990, there were few differences in these characteristics by racial/ethnic and income groups. Even as late as 1995–1996, the year before CSR, schools with high percentages of nonwhite and low-income students were slightly more likely than other schools to have inexperienced teachers who lacked full certification and postgraduate schooling. By 1999, large gaps in teacher qualifications emerged between schools attended by nonwhite and low-income students compared with other schools. For black students in schools with more than 75 percent of the students enrolled in subsidized lunch programs, nearly 25 percent had a first- or second-year teacher; almost 30 percent had a teacher who was not fully certified. At the other extreme, for white students attending schools with 25 percent or less of the students enrolled in subsidized lunch programs, only 12 percent had a first- or second-year teacher, and only 5 percent had a teacher who was not fully credentialed. These differences reflect the varying level of difficulty that many schools experienced in attempting to attract and retain teachers following the implementation of CSR.

Student Achievement

Evaluating the effect of CSR on student achievement is challenging for a number of reasons. First and probably most important, there are no statewide test scores in the years immediately preceding the implementation of CSR. Statewide tests began in 1997–1998, the second year of CSR. Thus, although much can be learned about the costs and benefits of CSR, its total effect on achievement will never be known.

Second, not all schools were able to participate in the program immediately, probably because of shortages of space and qualified teachers. Because participating schools had to reduce class size in first and second grade before reducing class size in other grades, adoption of CSR in the first and second grades was nearly complete by 1997–1998. Therefore, it is nearly impossible to compare achievement in 1997–1998 of schools that implemented CSR in these grades with ones that had not implemented CSR. However, nearly one-third of the schools had not implemented CSR in third grade as of 1997–1998. Consequently, it is possible to compare achievement in schools that reduced class sizes in third grade to achievement in those that had not.

The decision to implement CSR is not random but is based on financial or space constraints or a lack of qualified teachers. These factors also affect student achievement, making it extremely difficult to separate the effects of these constraints from the effects of CSR. The state-supported CSR Research Consortium has attempted to control for these differences between schools, but it is not clear that the consortium completely controlled for these between-school differences in its analysis of achievement.

Rather than looking at changes between schools, our analysis examines changes within schools in average class size in third grade between 1997–1998 and 1999–2000. We measure the effects of these changes, along with changes in teacher characteristics, on third-grade mathematics and reading achievement in California. This technique allows us to consider two effects of class size reduction on student achievement: the effects from the reduction in class size and the effects from the change in the teacher force. One main finding of Tennessee's Student/Teacher Achievement Ratio (STAR) experiment was that, all else equal, smaller classes are associated with higher achievement. In California, though, all else was not equal. As noted above, there were dramatic changes in teacher characteristics following the reduction of average class size, particularly for nonwhite students. This increase in the number of inexperienced and not fully certified teachers is to be expected when hiring a large number of new teachers. The effects of these increases should disappear as these new teachers acquire experience and full certification.

However, CSR likely had a more profound effect on the teacher workforce than simply increasing the number of inexperienced and uncertified teachers. Thousands of additional teaching positions were created, but thousands of additional teachers were not. Therefore, much of the increase in teachers consists of individuals who would not have been hired as teachers in the absence of CSR, especially given the availability of jobs with better pay and working conditions. If these teachers continue to be of lower quality than other teachers even after they have acquired additional experience, certification, or education, then CSR has the potential to create a long-term reduction in teacher quality.

The analysis shows that a ten-student reduction in class size (the average under CSR) raises the percentage of third-grade students who exceed the national median test score by roughly 4 percentage points in mathematics and 3 percentage points in reading. These findings are slightly larger than the effects found by the consortium in its analyses. Unlike the consortium, we find substantial variation by school in CSR's effect on achievement. Schools with more low-income students likely receive larger benefits, whereas schools in rural areas and those in which a high proportion of the students are black (primarily in Los Angeles Unified School District) appear to benefit little if at all from smaller classes.

The relationship between teacher characteristics and achievement is much weaker. The only indicator that is systematically linked to student achievement in third grade is experience. Having a new teacher reduces the percentage of students who exceed the national median by roughly 3 percentage points in both mathematics and reading. There is little or no evidence that teacher education or certification is significantly related to student achievement in third grade. However, the finding for certification could be influenced by the lower quality of the certification data.

One possible explanation for the variation in class size effects is the timing of CSR. As mentioned above, nearly one-third of schools had not implemented CSR in 1997–1998. The results show that for schools with a large percentage of black students, those recently implementing CSR have much smaller benefits from class size reduction than do other high-percentage black schools. A similar result is found for rural schools. In contrast, there is little evidence of a difference in class size benefits based on CSR timing in schools serving predominantly middle-class, nonblack students.

However, the hiring of new teachers explains only a portion of the difference in class size effects by CSR timing. Recent research in Texas and elsewhere has documented that hard-to-measure teacher attributes have an important effect on student achievement. It is likely that the smaller effects of class size for recent CSR implementers are related to changes in these hard-to-measure components of teacher quality, given that they are related to easier-to-measure teacher attributes such as experience. The schools that do not appear to benefit from CSR are the same schools that had trouble hiring experienced, certified teachers before CSR.

The estimated decline in third-grade teacher quality in this analysis probably understates the actual quality decline that accompanied the implementation of CSR statewide in two ways. First, the effects of class size reduction on teacher quality extend beyond the grades where class sizes were reduced. An analysis of fifth-grade achievement shows that class size reduction in third grade is negatively associated with achievement in fifth grade for schools with a high percentage of black students. Such a finding is consistent with the story that the movement of many fourth- and fifth-grade teachers into the early grades meant that schools had to rapidly expand hiring in all grades, not just K–3. Second, and perhaps even more important, the available data are not able to capture between-school changes in instructional effectiveness. Yet, it is widely believed that many teachers switched schools as a result of class size reduction. Schools must fill not only the additional positions created by their own efforts to reduce class size but also positions vacated by the departure of teachers for newly created opportunities at other schools.

Policy Implications

Because it is difficult to calculate the magnitude of the benefits of CSR in terms of higher student achievement, the question of whether money would have been better spent on other aspects of schools such as higher teacher salaries, expanded and improved pre-school, technology, or other programs is very hard to answer. Nevertheless, there is clear evidence that, controlling for changes in teacher quality, smaller classes raised student achievement and the effects were larger in schools serving predominantly lower-income students. Unfortunately, these schools tended to suffer the largest deterioration in teacher quality as measured by experience and certification. A better approach to class size reduction would have been to reduce class sizes in a subset of schools each year, starting with low-performing schools serving high-poverty populations.

This would have limited the departure of teachers for newly created jobs in suburban schools, lessened the overall competition for new teachers, and reduced inequality in academic performance.

The results concerning new teachers and their struggles highlight the importance of policies targeting new teachers, such as the Beginning Teacher Support Act (BTSA). These types of programs have two potential benefits. First, they can minimize the adverse effects of new teachers by helping them adapt to the classroom more quickly and effectively. Making these teachers more productive and effective also reduces the stress of the job, thereby reducing teacher turnover. Effective programs that assist new teachers will become even more essential in the near future, as enrollment and teacher retirements both increase. Future research should carefully analyze the effectiveness of programs such as BTSA on student achievement.

Although we find that experience matters, the relationship between certification and achievement is much less clear. Our results show that California's certification system in the late 1990s had little if any relationship to student achievement, suggesting that policies that prevent uncertified teachers from teaching are unlikely to raise student achievement. The concern about the weak relationship between certification and student achievement is well known to California policymakers, and the California Commission on Teacher Credentialing (CCTC) is currently reforming the certification process. It is hoped that the effect of these reforms will be studied rigorously.

At the same time, the available data in California on certification limit the inferences that can be drawn. For example, it is not possible to distinguish between the two types of full credentials (preliminary versus professional clear). Nor can the data from the CCTC be linked to the California Basic Educational Data System (CBEDS) from the Department of Education. It is not possible to match individual student test scores from one year to the next to measure student growth. Thus, several improvements to the current data collection system are needed to conduct an in-depth evaluation of the effects of certification on student achievement.

Contents

Figures

Tables

Acknowlegments

We would like to thank numerous people who assisted us with this research project. Pedro Cerdan and Jacklin Chou provided excellent research assistance, and Anne Danenberg gave us ongoing help in the construction of the datasets. Julian Betts, Dominic Brewer, Anne Danenberg, Norton Grubb, and Belinda Reyes provided thoughtful reviews of preliminary drafts, and seminar participants at PPIC also offered very helpful comments along the way. Gary Bjork, Susanna Cooper, Joyce Peterson, and Patricia Bedrosian provided valuable editorial assistance. Finally, we are extremely grateful to Mark Baldassare, Gary Hart, Kim Rueben, and Jon Sonstelie for their invaluable assistance and support throughout the project. The authors are solely responsible for any errors of fact or interpretation.

1. Introduction

In response to widespread dissatisfaction with the public schools, California has implemented a number of education reforms over the past decade. The legislature and governor enacted the nation's largest class size reduction (CSR) law with overwhelming support in the summer of 1996.[1] Educators and policymakers hoped that this policy would lead to large gains in student achievement, echoing gains attributed to smaller class sizes in the Tennessee Student/Teacher Achievement Ratio (STAR) experiment.[2]

The goal of the CSR program was to reduce class sizes to no more than 20 students in each grade from kindergarten through third grade. In the initial year of the program, 1996–1997, school districts received $650 for each K–3 student in a full-day CSR program and $325 for each student in a half-day CSR program (i.e., half the instructional minutes are spent in classes of 20 children or fewer). This amount is increased every year by a cost of living allowance. In 1996–1997, districts also received a one-time amount of $25,000 for newly created classrooms. That amount rose to $40,000 in later years. In 1996–1997 and 1997–1998, schools could use classrooms that were smaller than the state minimum of 960 square feet for grades 1–3 and 1,350 square feet for kindergarten (Bohrnstedt and Stecher, 1999). However, by 1998–1999 class sizes for CSR classrooms had to be the same size as other classrooms in the school. Overall, CSR reduced the state average class size for K–3 from approximately 29 students in 1995–1996 to 19 in 1999–2000.

---

[1]The information on CSR comes from the California Department of Education's "Fingertip Facts" on Class Size Reduction (http://www.cde.ca.gov/classsize/facts.htm). For more information, also see Bohrnstedt and Stecher (1999).

[2]Note that the Tennessee STAR experiment has nothing to do with California's Standardized Testing and Reporting system even though both programs share the STAR acronym. Throughout this report, STAR refers to the Tennessee STAR experiment.

Districts were free to choose the schools in which to reduce class sizes, although they were not free to choose the grades within a school. Each participating school had to reduce class size in first grade before reducing class size in second grade. Similarly, each school had to reduce class size in second grade before reducing class size in kindergarten or third grade. Almost every school district participated in the initial year, although not in every school and grade. In the second year of CSR, 1997–1998, over 95 percent of the first- and second-grade classrooms qualified for CSR, compared with only 70 percent for kindergarten and third-grade classrooms (Bohrnstedt and Stecher, 2000). Participation in kindergarten and third grade increased to roughly 85 percent in 1998–1999.

There are few if any differences in average class size in the affected grades (K–3) by income level or race/ethnicity at any point during the decade, as shown in Appendix Table B.1.[3] Low-income and nonwhite students on average do not attend larger classes than other students. In fact, additional analyses of average K–3 class size found little variation between urban, suburban, and rural schools. Appendix Table B.2 shows little variation in average class size for grades four and five.

One of the biggest challenges to districts implementing CSR was the need to hire more teachers. Extra classes created a need for 25,000 additional teachers statewide. In contrast, fewer than 4,000 new teachers were hired in kindergarten through third grade in the year before CSR (1995–1996). Some districts were already beset by staffing difficulties before class size reduction, and the need to hire many additional teachers exacerbated the problem. Ross (1999) describes the influx of inexperienced, noncertified teachers into elementary schools in South Central Los Angeles following CSR. This influx, he reports, was prompted in part by the departure of experienced teachers to newly created positions in more affluent communities.

---

[3]Because we are interested in the typical classroom, special education and other "alternative" classrooms are excluded from the analysis. This exclusion eliminates the possibility of confounding class size differences with differences in the number of special education classes, which are typically smaller than regular classes and more numerous in low-income schools and in schools with more nonwhite students.

This influx of new teachers sets up the possibility of a serious side effect of CSR: The need to hire large numbers of additional teachers may have reduced teacher quality and offset some or most of the benefits of smaller classes. A crucial question is whether any decline in quality is a temporary consequence of the need to hire large numbers of inexperienced teachers or a longer-term result of employing less-qualified applicants who would not have been considered before the enactment of CSR. A distressing possibility is that any decline in teacher quality may have been concentrated in schools serving predominantly lower-income and minority students, which have the most difficulties attracting and retaining highly skilled teachers.

Many organizations have studied the effects of CSR on a variety of outcomes, including achievement. Because of the absence of statewide testing before 1997–1998 and other aspects of policy implementation, it will not be possible to ever know the full effect of CSR on the distribution of teacher quality and student achievement. Moreover, because most research, including the widely publicized Tennessee STAR experiment, investigates the effect of smaller classes *holding all other factors including teacher quality constant*, the results provide information strictly on the benefits of smaller classes. Differences between California and Tennessee or other research sites may limit the relevance of previous work on class size even more.

Despite the impediments to understanding the full effect of CSR, it is still quite important to learn as much as possible about how it is affecting California schools. The state commissioned the CSR Research Consortium to study the effects of CSR on student achievement and other outcomes. By comparing schools that implemented CSR with those that did not, the first two reports by the consortium found small but positive effects of CSR implementation on achievement in second and third grade. They found that these benefits were not concentrated in any particular demographic group such as low-income or nonwhite students; all students benefited equally. However, there were two problematic assumptions in these reports. First, they assumed that the effect of CSR was the same for each school: Schools that reduced class size from 21 to 20 received the same benefit as schools that reduced class size from 31 to 20. Second, the consortium's analysis of achievement in third grade also explicitly assumed that CSR had no effect on achievement in grades four and five, despite its own findings that the ancillary effects of CSR increased the number of new teachers in fourth and fifth grade by as much, if not more, than the number of new teachers in kindergarten through third grade.

The consortium's third report used state-level data to look at the effects of CSR implementation in kindergarten through third grade, rather than focusing on CSR in second and third grade. The report found no systematic relationship between CSR and achievement, but the dataset used for this analysis was extremely small. The fourth and final consortium report, due in the summer of 2002, will use much more extensive student- and school-level data.

The achievement results from all three consortium reports contradicted the anecdotal stories given by principals and superintendents describing the struggles faced by schools serving low-income and nonwhite students in their attempts to benefit from CSR. The Center for the Future of Teaching and Learning and the consortium both documented the staffing struggles of these schools following CSR. Although the center did not study achievement, it suggested that these staffing inequities are likely to lead to differences in achievement.

The primary objective of our report is to determine the mechanisms by which CSR affected student achievement. Rather than simply asking, "Does CSR raise achievement?" as the consortium did, we also ask, "*How* does CSR raise achievement?" in an attempt to learn as much as possible about changes in teacher quality that may have accompanied CSR.

In addition to an analysis of all students combined, we also consider differences by income, race/ethnicity, community type, and even specific school district. Previous work in California and across the nation has shown that schools serving lower-income students in urban districts face some of the greatest obstacles to raising achievement, in large part because of the difficulties of attracting and retaining high-quality teachers. By highlighting the experiences of the state's most disadvantaged students, we provide the detailed information for state policymakers to consider in their efforts to improve the quality of California's public schools.

To examine the effect of class size reduction on test scores, we focus on mathematics and reading achievement in third grade. Our analysis uses data from the 1997–1998 and 1999–2000 school years to examine changes in class size within schools.[4] We focus on third grade because CSR was not implemented in a substantial number of third-grade classrooms by 1997–1998.[5] During our time period, some schools had large changes in class sizes and teacher characteristics because of CSR implementation, whereas other schools had much smaller changes as a result of other factors such as enrollment. We consider the effect of smaller classes as well as the effect of the drastic changes in the teacher workforce. We also investigate whether the relationship between class size and achievement depends on the timing of CSR.

The report is structured as follows. Chapter 2 briefly describes public school enrollment patterns to provide the context for our analysis. Chapter 3 describes teacher characteristics (i.e., experience, certification, and education) before and after the implementation of CSR, especially the distributions of teacher characteristics by community type, student racial/ethnic composition, and income. Following the description of trends in teacher characteristics, Chapter 4 analyzes the effects of class size reduction on mathematics and reading achievement in third grade. The final chapter summarizes the findings and offers a number of policy implications based on the empirical analysis. Additional technical materials appear in appendices at the end of the report.

---

[4]Unfortunately, consistent test score data were not available for the pre-CSR period.

[5]Standardized tests were not given in kindergarten, where implementation also was not universal.

2. Elementary School Demographic Composition

Introduction

Demographic composition may affect academic achievement in a number of ways, including the influences of peers, structure of the curriculum, and instructional quality. There remains little consensus on the magnitude or even direction of peer group influences, and we do not address that issue in this work. Rather, we focus on the link between student demographic composition and the quality of instruction. Specifically, an important aspect of this report is the extent to which any changes in instructional quality following the implementation of CSR varied systematically by racial/ethnic or income characteristics of the students. Research shows that schools with high shares of black, Hispanic, or low-income students tend to have a more difficult time attracting and retaining teachers, although there is little evidence on the underlying causes of the association between teacher preferences and school demographic composition.[1] Some teachers may prefer specific student characteristics, but a much more likely explanation is that the link between teacher preferences and student demographic composition is driven by factors including the quality of facilities, safety, distance to work, and other aspects of working conditions.

This chapter provides background information on trends in California elementary school demographic composition during the 1990s. Following a description of statewide enrollment shares, we examine enrollment patterns for urban, suburban, and rural districts. To learn even more about the large urban districts that appear to have the most difficulty attracting and retaining teachers (see Chapter 3), we also

[1]See Lankford, Loeb, and Wyckoff (2001) and Hanushek, Rivkin, and Kain (1999).

document enrollment patterns for each of the six largest districts in the state.

Statewide Trends

California has always been a destination for immigrants from other states and countries, and the past ten years have certainly been no exception. The large-scale immigration of Hispanics and Asians into California during the 1980s and 1990s has had a dramatic effect on school demographic composition. Figure 2.1 shows that the share of all public school students who are Hispanic increased from 36.2 percent to 46.3 percent during the 1990s; Hispanics now constitute a plurality in California public schools. Asian and black enrollment remained at roughly 10 percent and 9 percent, respectively, throughout the decade, and the increase in Hispanic enrollment was offset by a decline in the white (non-Hispanic) enrollment share, from 44.4 percent in 1990–1991 to 34.4 percent in 1999–2000. In absolute terms, white enrollment fell by more than 150,000 students (over 15 percent) at a time when total enrollment rose by roughly 11 percent.

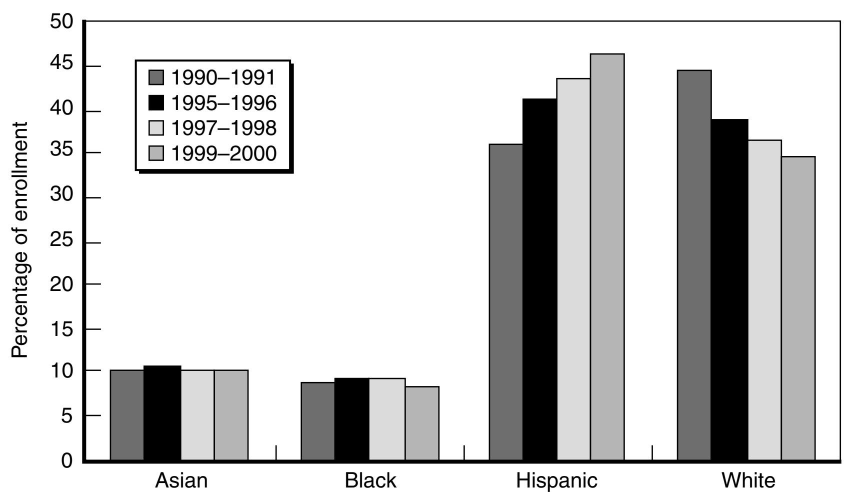


Figure 2.1—Percentage Enrollment in Elementary Schools, by Race/Ethnicity, 1990–2000

The categories of Hispanic and Asian comprise students from numerous countries of origin, including both recent immigrants and those whose families have lived in the United States for many generations. Because recent immigrants are more likely to lack English proficiency and to live in lower-income families, it is important to go beyond the broad classifications of Hispanic and Asian. In an attempt to distinguish newly arrived immigrants, who are less likely to be proficient in English and more likely to come from low-income families, Table 2.1 reports the distribution of Hispanic and Asian elementary school students, showing the proportion of English learners (ELs) and non-English learners during the 1990s.[2] This separation is a crude division both because nativity and English proficiency are not perfectly correlated and because classification criteria for limited English proficiency varies among schools and districts as well as over time. Nevertheless, the results show that the percentage of Hispanics not proficient in English increased by more than 10 percent, whereas the corresponding percentage for Asians declined by a similar proportion. In 1999–2000, roughly 60 percent of Hispanics and 38 percent of Asians in elementary school were classified as English learners.

An alternative to dividing students by race/ethnicity is division by family income. The available income measure for students in California is the school's percentage of students enrolled in free or reduced-price

Table 2.1

Percentage Enrollment of Asian and Hispanic Students, by English Language Proficiency, 1990–2000

	1990–1991	1995–1996	1997–1998	1999–2000
Hispanic				
EL	53.5	59.7	59.9	59.2
Non EL	46.5	40.3	40.1	40.8
Asian				
EL	42.7	44.8	41.9	38.0
Non EL	57.3	55.2	58.1	62.0

[2]EL students were formerly known as limited English proficient or LEP students.

lunch programs (students are eligible to receive a subsidized lunch if their family income is less than or equal to 185 percent of the poverty level). Figure 2.2 depicts the enrollment pattern from 1990–1991 to 1999–2000 for four categories of schools, based on the percentage of students enrolled in subsidized lunch programs: 0 to 25 percent, 25 to 50 percent, 50 to 75 percent, and 75 to 100 percent.[3] The figure illustrates the dramatic growth over the decade in the enrollment of low-income students. In 1990–1991, 33 percent of the students were in low-poverty schools (0 to 25 percent subsidized lunch), compared to 23 percent in 1999–2000. Conversely, the enrollment share for high-poverty schools (75 to 100 percent subsidized lunch) increased from 20 percent in 1990–1991 to 36 percent in 1999–2000. The bulk of the increase of low-income students occurred between 1990–1991 and 1995–1996.

Given the changes in racial/ethnic composition and the fact that recent Asian and Hispanic immigrants are more likely than U.S.-born

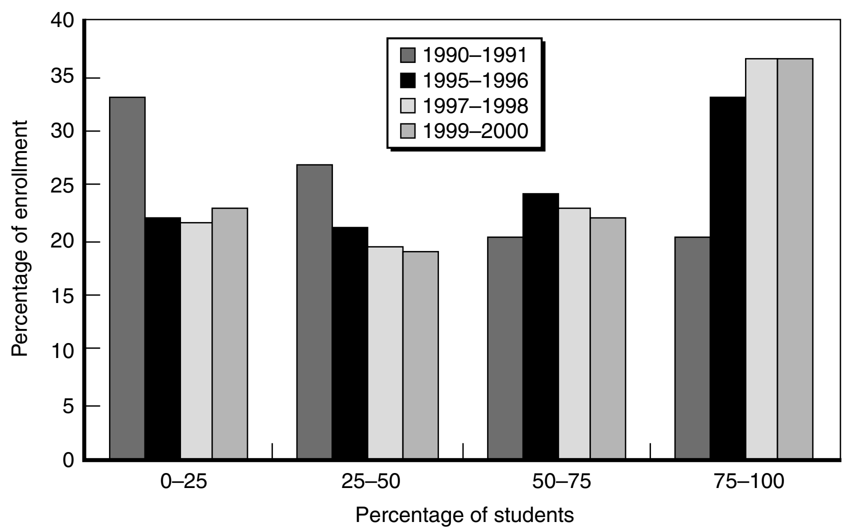


Figure 2.2—Percentage Enrollment in Free or Reduced-Price Lunch Programs, 1990–2000

[3]For comparison purposes, Appendix Table B.3 presents the distribution of the number of schools in each subsidized lunch category, from 1990–1991 to 1999–2000. The table also contains the distribution for urban schools.

whites to be poor, the trends for income are not surprising. Perhaps a more important question is how the percentage of students enrolled in free or reduced-price lunch programs differs across racial/ethnic groups today. To the extent that income and race/ethnicity have independent links with teacher labor supply, it is important to identify differences along each of these dimensions.

To explore the relationship between race/ethnicity and the percentage of students enrolled in free or reduced-price lunch programs, we calculated separate distributions for Asians, blacks, Hispanics, and whites for the 1999–2000 school year. Figure 2.3 clearly illustrates that the concentration of poverty in the schools varies dramatically by race/ethnicity. For Hispanic and black students, the enrollment percentage increases monotonically with the percentage enrolled in free or reduced-price lunch programs. Over 56 percent of the Hispanic students are in high-poverty schools, and only 6.5 percent are in low-poverty schools. The percentages for black students show a similar pattern: 45.3 percent in high-poverty schools compared to 9.7 percent in low-poverty schools. White students, on the other hand, are primarily concentrated in low-poverty schools (44.4 percent). Only 10.1 percent

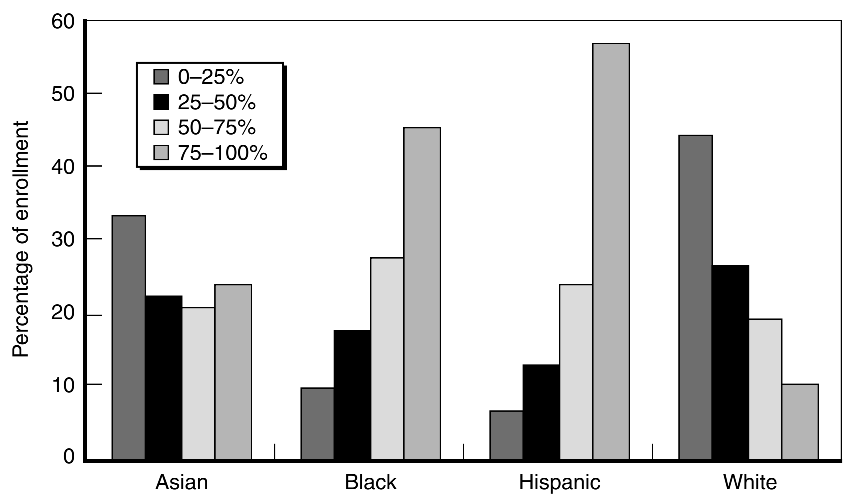


Figure 2.3—Percentage Enrollment in Free or Reduced-Price Lunch Programs, by Race/Ethnicity, 1999–2000

of white students are in high-poverty schools. Finally, Asian students are more evenly distributed across the school income distribution (33.1 percent in low-poverty schools and 23.8 percent in high poverty schools). Note that these patterns are relatively constant throughout the decade (see Appendix Figure B.1).

Differences by Community Type

The statewide figures provide a broad overview of elementary school enrollment patterns, but it is also important to examine systematic differences among districts. Table 2.2 describes separate school enrollment patterns for urban, suburban, and rural schools (as defined by the California Department of Education) in California during the 1990s. As expected, there are large differences by ethnicity in the percentages of students in the various community types. Throughout the decade, blacks and Hispanics were much more likely to attend school in urban districts, whereas whites were much more likely to attend school in suburban or rural districts.

Figure 2.4 illustrates the distribution of students in 1999–2000 by race/ethnicity and percentage of students enrolled in free or reduced-price lunch programs for each community type. Across all racial/ethnic groups, suburban students are much less likely to attend high-poverty schools than rural or urban students. This difference across community

Table 2.2

Percentage School Enrollment, by Race/Ethnicity and Community Type

	Urban	Suburban	Rural
1990–1991			
Asian	12.9	11.4	3.2
Black	13.8	6.5	2.6
Hispanic	45.4	26.6	30.1
White	27.4	54.9	62.4
1999–2000			
Asian	11.7	13.3	3.9
Black	13.4	6.4	2.9
Hispanic	53.5	34.1	39.2
White	20.8	45.5	52.0

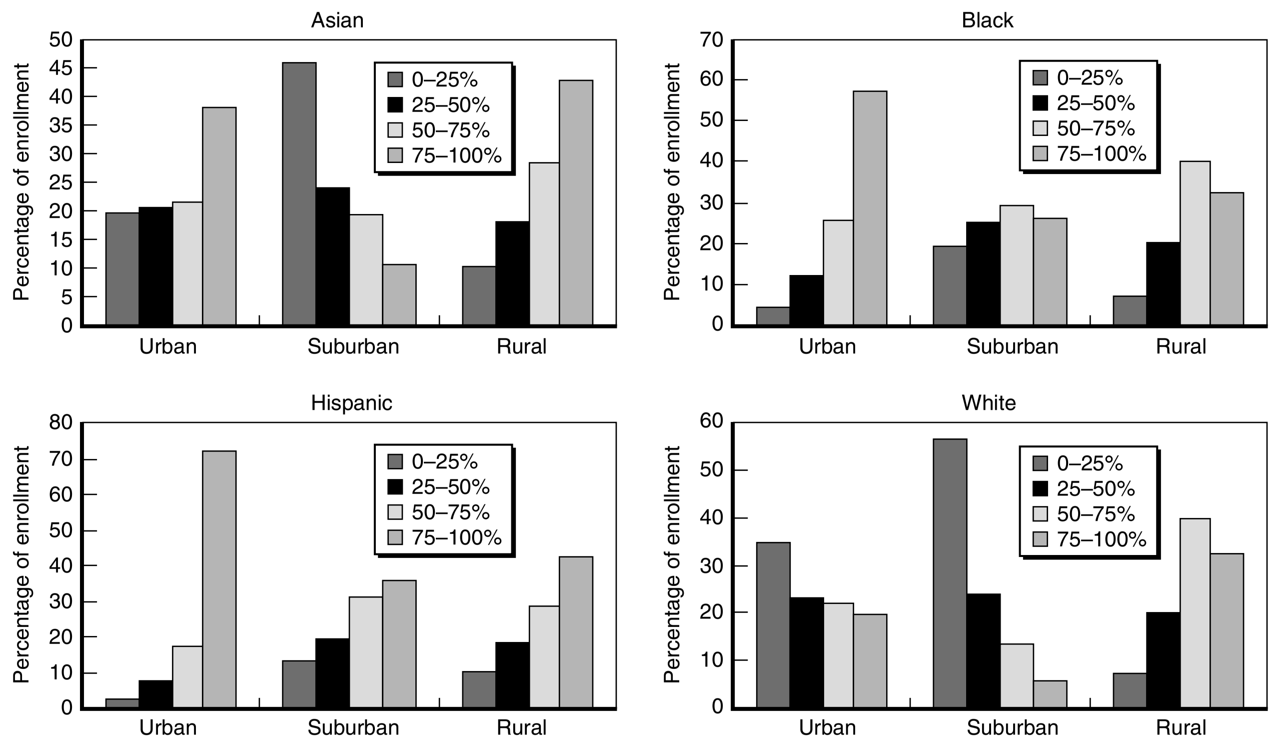


Figure 2.4—Percentage Enrollment in Free or Reduced-Price Lunch Programs, by Race/Ethnicity, Percentage of Students Enrolled in Free or Reduced-Price Lunch Programs, and Community Type

types is most pronounced for Asian and white students, where most suburban students attend low-poverty schools. The enrollment percentage among Hispanic students increases with the percentage enrolled in free or reduced-price lunch programs in all three community types; over 70 percent of urban Hispanic students are in high-poverty schools. Urban and rural black students are much more likely to attend high-poverty schools, whereas suburban blacks are likely to attend low-, middle-, and high-poverty schools in roughly the same proportions.

Enrollment Patterns in the Six Largest School Districts

The aggregate trends and differences by community type are quite informative, yet they may conceal substantial differences across districts. Because large urban districts tend to have high fractions of poor, nonwhite students, this report focuses additional attention on such districts. Historically, the six largest school districts in California have been Fresno, Long Beach, Los Angeles, Oakland, San Diego, and San Francisco.[4] These six districts represent 20 percent of the enrollment in California elementary schools and 50 percent of the enrollment in urban schools.

There is substantial variation among the six districts in demographic composition, although each experienced qualitatively similar demographic shifts during the 1990s. Table 2.3 contains the enrollment percentages by year, race/ethnicity, and district. Each district has one dominant racial/ethnic group. In four of the districts, it is Hispanics; in Oakland it is blacks; in San Francisco it is Asians.

Nevertheless, Hispanic enrollment as a percentage of the total increased in all six districts, so that by 1999–2000 the percentage of Hispanic students ranged from 25 percent in San Francisco to 73 percent in Los Angeles. The increase in percentage Hispanic was most dramatic in Oakland, nearly doubling between 1990–1991 and 1999–2000. During this same period, white enrollment dropped in all six districts and did not exceed one-quarter of the enrollment in any of the districts

[4]Santa Ana is now the sixth largest school district in the state because of recent enrollment growth. Oakland is currently the seventh largest.

Table 2.3

Percentage Enrollment in the Six Largest School Districts, by Race/Ethnicity, 1990–2000

	1990–1991	1995–1996	1997–1998	1999–2000
		Asian		
Fresno	22.4	23.6	20.3	18.2
Long Beach	22.4	18.6	16.2	14.5
Los Angeles	6.9	6.0	5.8	5.6
Oakland	18.4	19.2	17.2	16.0
San Diego	17.8	17.9	17.1	16.7
San Francisco	43.3	45.6	46.9	47.7
		Black		
Fresno	9.8	10.5	11.5	11.3
Long Beach	18.2	21.4	20.4	18.9
Los Angeles	14.9	13.6	13.1	12.2
Oakland	55.4	51.4	51.0	48.5
San Diego	16.5	17.1	17.2	16.4
San Francisco	20.2	19.0	17.2	16.8
		Hispanic		
Fresno	36.9	43.8	47.9	51.4
Long Beach	33.4	40.7	46.6	50.6
Los Angeles	65.8	69.8	71.1	72.9
Oakland	16.9	22.3	25.8	29.6
San Diego	28.8	35.8	38.7	41.1
San Francisco	21.7	22.2	24.2	24.8
		White		
Fresno	30.9	22.1	20.3	19.1
Long Beach	26.0	19.4	16.8	15.9
Los Angeles	12.4	10.7	10.0	9.3
Oakland	9.4	7.0	6.0	5.9
San Diego	36.8	29.2	27.1	25.7
San Francisco	14.8	13.3	11.7	10.7

(San Diego had the highest white enrollment share in 1999–2000, at 25.7 percent). The Asian enrollment share also declined in all large districts between 1990–1991 and 1999–2000, with the exception of San Francisco.

Figure 2.5 looks at the relationship between race/ethnicity and income in 1999–2000. The figure shows the percentage of students of each race/ethnicity who were in schools with 75 to 100 percent of the students enrolled in free or reduced-price lunch programs. There is substantial variation even among the six largest districts in the incidence

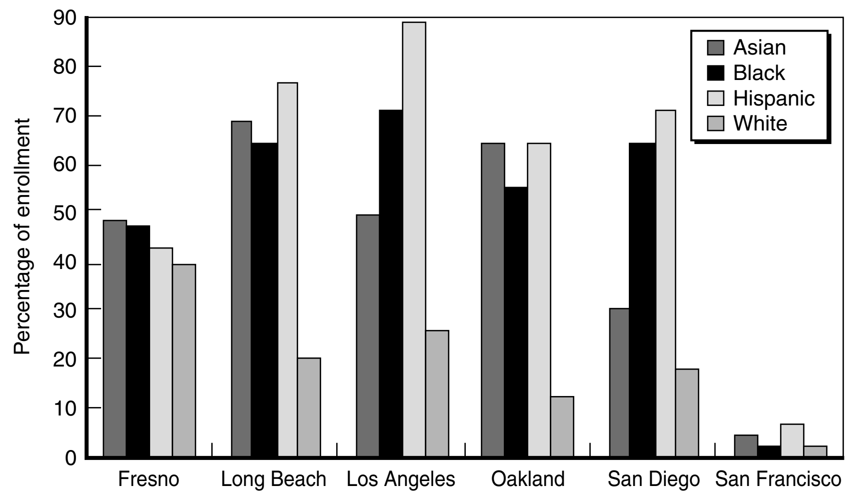


Figure 2.5—Percentage Enrollment in Low-Income Schools in the Six Largest School Districts, by Race/Ethnicity, 1999–2000

of high-poverty schools. Far smaller shares of Asians, blacks, Hispanics, and whites attend high-poverty schools in San Francisco than in other districts. There are also pronounced differences in the racial/ethnic poverty gaps within districts. The likelihood of attending a high-poverty school in Fresno is quite similar for all groups, whereas in the other districts nonwhites are much more likely to attend such a school. Hispanics are the most likely to attend high-poverty schools in five of the six districts, and the concentration of recent immigrant Hispanics in high-poverty schools is probably far higher than that for Hispanics as a whole.

Summary

Immigration and other factors led to substantial demographic changes in California public elementary schools during the 1990s. No single ethnic group constitutes a majority of students, and there are substantial differences in school demographic composition across the state. Such differences, even among the six largest districts, raise the possibility that class size reduction may have had dramatically different effects across the state. This depends largely on the strength of the relationship between teacher availability and student demographic composition, an issue to which we now turn.

3. Trends in Teacher Characteristics During the 1990s

Introduction

Changes in school enrollment patterns not only alter peer group composition but also potentially affect a variety of other aspects of school quality, including the quality of instruction. Schools with high percentages of nonwhite or low-income students appear to face more staffing difficulties than do other schools. The observed pattern of teacher transitions out of high-percentage nonwhite schools and the difficulty many such schools experience in trying to attract and retain certified teachers suggest that average instructional quality is likely to be lower in high-percentage nonwhite schools.[1] Consequently, the creation of teaching jobs by the CSR legislation was likely to exacerbate staffing difficulties in schools that were struggling to attract and retain teachers before the implementation of CSR.

A growing body of research documents the preeminent role of teacher quality in the determination of educational outcomes, particularly compared with class size and other observable school inputs.[2] Identifying whether a teacher is effective is not an easy task, but identifying the attributes that make a teacher effective is even more difficult. The attributes of high-quality teachers are not entirely captured by observable characteristics such as education and experience. Nevertheless, certain teacher characteristics appear to affect student

---

[1]See Hanushek, Rivkin, and Kain (1999) for evidence on teacher transitions in Texas. Reichardt (2000) provides preliminary evidence on teacher transitions in California.

[2]See Sanders and Horn (1994) and Rivkin, Hanushek, and Kain (2000).

achievement. Whereas average teacher experience is not closely linked to student achievement, recent work suggests that first- and second-year teachers perform markedly worse than more experienced colleagues.[3] Limited evidence shows that teachers lacking certification and training in a given field such as mathematics are likely to perform worse than other teachers.[4] Other potential indicators of quality are the test scores of teachers and the quality of the college they attend.[5]

Moreover, because schools use education, certification, experience, and other observable characteristics in judging potential teachers, their distributions provide evidence on the difficulties schools (and districts) experience in teacher labor markets. For example, the state provides strong incentives for the hiring of certified rather than uncertified teachers, so that a high rate of uncertified teachers indicates that many schools are having serious difficulty attracting and retaining teachers. Such difficulties almost certainly lower the quality of instruction.

This chapter documents changes in the distribution of teacher characteristics in California public elementary schools during the 1990s, highlighting the apparent effects of CSR.[6] Similar to the previous chapter, the analysis begins with statewide trends in teacher characteristics. Subsequently, changes for all urban districts combined and for each of the six largest districts are described. The tables report teacher characteristics for the school years 1990–1991, 1995–1996, 1997–1998, and 1999–2000. The year 1995–1996 is not only the midpoint of the decade but also the year before the implementation of CSR.

The description of the distribution of teacher characteristics in this report builds upon previous work. The reports by the CSR Research Consortium (Bohrnstedt and Stecher, 1999, 2000, 2002) describe the increase in both the use of emergency certification and the percentage of

---

[3]See Rivkin, Hanushek, and Kain (2000).

[4]See Fetler (1999) and Goldhaber and Brewer (2000).

[5]See Ehrenberg and Brewer (1994) for evidence on college selectivity and Ferguson (1998) for evidence on teacher test scores.

[6]Because the trends in teacher characteristics are similar for the grades reducing class size (K–3) and other grades (4–5), we combine the results for all elementary grades.

beginning teachers following the implementation of CSR. They also document differences by student income and race/ethnicity in the levels and changes in these teacher characteristics. Betts, Rueben, and Danenberg (2000) present a comprehensive analysis of differences in teacher characteristics both within and between public school districts in California for the single school year 1997–1998.[7] In particular, they highlight the substantial differences in teacher characteristics within districts despite the presence of districtwide salary schedules. These results strongly suggest that California teachers place a great deal of weight on student characteristics and other determinants of working conditions in deciding where to work.

Although our description of California elementary school teachers covers much of the same ground, there are several key distinctions from this earlier work. First, rather than combining all teachers with less than three years of experience into a single category, we create separate categories for first- and second-year teachers and group all remaining teachers with at least two years of experience into a single category. Second, the studies mentioned above examined differences by community type, student racial/ethnic composition, and student income, but they did not consider the joint influences of these three sets of factors. The results in the previous chapter illustrate that the distribution of students within each income category differs by race, and these differences would be overlooked if focusing on either race/ethnicity or income categories.[8] Therefore, our statewide analysis considers the joint distribution of race/ethnicity and enrollment in free or reduced-price lunch programs. Within each community type, there is a high correlation between race/ethnicity and income, so we consider the joint distribution of community type and race/ethnicity.

Experience

Table 3.1 describes the percentages of new and second-year teachers in an elementary school for the 1990s, not just the period immediately

---

[7]They look at middle and high schools, as well as elementary schools.

[8]As in the previous chapter, the measure of income is the percentage of students enrolled in free or reduced-price lunch programs.

Table 3.1

Teacher Experience, by Student Race/Ethnicity and Percentage of Students Enrolled in Free or Reduced-Price Lunch Programs, All Students

	% with 0 Years Experience				% with 1 Year Experience			
% of Students in Lunch Program	1990–1991	1995–1996	1997–1998	1999–2000	1990–1991	1995–1996	1997–1998	1999–2000
Asian								
0–25	4.7	4.6	9.8	6.6	5.4	4.4	8.7	6.7
25–50	5.5	5.1	10.3	7.2	6.0	4.5	9.3	7.9
50–75	7.2	5.1	12.2	8.4	6.0	5.2	9.8	9.7
75–100	6.8	6.4	13.3	9.6	7.3	6.5	10.4	11.0
All	5.8	5.3	11.4	7.8	6.0	5.2	9.5	8.6
Black								
0–25	5.1	4.2	8.8	6.4	6.5	4.0	8.4	6.7
25–50	6.4	5.2	9.6	7.3	6.1	4.6	9.0	7.5
50–75	9.1	5.9	11.9	7.8	6.9	5.6	9.5	9.8
75–100	8.6	9.0	15.7	11.9	6.8	7.9	11.1	13.1
All	7.6	7.0	13.1	9.5	6.6	6.3	10.1	10.6
Hispanic								
0–25	5.8	4.6	9.0	6.6	6.5	4.1	8.9	7.0
25–50	6.4	5.5	10.1	6.8	6.4	4.8	9.3	7.1
50–75	8.0	6.3	12.7	8.1	7.3	5.6	9.9	9.4
75–100	7.6	7.6	14.9	10.4	7.0	6.8	10.2	11.5
All	7.2	6.8	13.4	9.2	6.9	6.0	9.9	10.2
White								
0–25	4.9	4.3	8.5	6.0	5.9	4.4	8.7	6.6
25–50	6.2	4.7	8.8	5.7	6.4	4.2	8.8	6.3
50–75	7.3	5.7	11.4	6.7	7.0	4.8	9.3	7.7
75–100	8.4	7.0	12.1	8.3	6.7	6.0	10.3	9.0
All	5.7	4.9	9.6	6.3	6.2	4.6	9.0	7.0
All Students								
	6.4	5.9	11.8	8.0	6.5	5.4	9.6	8.9

NOTE: The percentages are weighted by the number of students in each racial/ethnic category.

surrounding CSR. Because the patterns for experience, certification, and education are similar for grades K–3 and grades 4–5, the findings are presented for all elementary grades combined. The tables in this chapter—including Table 3.1—are constructed in the following way. As in Chapter 2, schools are divided into four categories according to the percentage of students enrolled in free or reduced-price lunch programs. Then, the statistic, such as the average of each school's percentage of new teachers, is calculated from all the schools in that lunch category. In the case of the final row ("all students"), schools are weighted according to total enrollment, and the reported percentage reflects the average for all students in that category. In the other rows, the calculations weight each school by the number of students in the specific ethnic group (e.g., Asians in the case of the first five rows). Therefore, these four rows report the averages for students in each of the four ethnic groups.[9] Differences among rows across ethnic groups within each lunch category provide information on the degree to which the distributions of class sizes or teacher characteristics differ by race/ethnicity independent of income.

For teacher experience, a clear pattern, based on percentage of students enrolled in subsidized lunch programs by 1995–1996, emerges: For all races/ethnicities, the percentage of teachers without experience or with one year of experience increases as the percentage of students enrolled in subsidized lunch programs increases. Moreover, the gap in teacher experience between low-poverty and high-poverty schools widens for nonwhites following the implementation of CSR. For students in the highest poverty schools (75 percent of students or more enrolled in subsidized lunch programs), the increase between 1995–1996 and 1997–1998 in the number of teachers with no experience ranges from over 70 percent for blacks to over 100 percent for Asians. Although the percentage of teachers in their first year declines for all demographic groups between 1997–1998 and 1999–2000, the rate in 1999–2000 still exceeds that in 1990–1991 for nonwhites in high poverty schools. An important question is whether the higher rates of novice teachers in high-poverty schools will persist into the future.

The large discrepancies in the percentage of new teachers by race/ethnicity and percentage of students enrolled in subsidized lunch programs, coupled with the similarity across these groups in the average reductions in class size (see Appendix Table B.1), are consistent with a substantial teacher movement away from high-poverty, high-percentage nonwhite schools. Reichardt (2000) finds exactly this type of teacher

---

[9]The calculations assume that there is no systematic variation within schools by race or ethnicity in the probability of having inexperienced or uncertified teachers.

movement within districts during the first year following CSR implementation. However, more complete data regarding teacher transitions among schools are not available for subsequent years.

Certification

The certification (or credentialing) process for elementary teachers in California involves multiple steps. Currently, there are two types of fully certified teachers.[10] A professional clear certification has the most requirements, although legislation passed in 1998 (SB 2042) attempts to streamline the certification process. For example, the required fifth year of study (30 semester units beyond the completion of a bachelor's degree) can now be satisfied through completion of either a blended four-year university program or an induction program such as the Beginning Teacher Support Act (BTSA). The requirements for the preliminary full certification, which are a prerequisite for the professional clear credential, include the completion of the following: a bachelor's degree (or higher)—including the completion of an approved teacher training program, the California Basic Educational Skills Test (CBEST), and a program or examination demonstrating subject matter competence. Teachers who possess either the professional clear or the preliminary certification are considered fully certified, and the California Department of Education data we use do not distinguish between these two types of certification.

Teachers who do not meet these qualifications are considered not fully certified—or uncertified—teachers. There are several types of uncertified teachers. University interns are teachers who are in the process of completing the required coursework for the teacher training program at a participating university. District interns have completed their bachelor's degree but have not satisfied the requirements for the teacher training program. Teachers who complete either intern program become preliminary certified teachers.[11]

---

[10]The information on certification comes from the California Commission on Teacher Credentialing website, www.ctc.ca.gov.

[11]Before 1997–1998, these teachers were called trainees.

If a school demonstrates that it is not able to fill all its teaching positions with certified or intern teachers, it may apply for permission to hire other uncertified teachers. There are two avenues for hiring such teachers: emergency permits and waivers.[12] Requirements for emergency teachers are a bachelor's degree, CBEST completion, and subject matter competence, through either an examination or coursework. The requirements for a waiver teacher are not specified, but waiver teachers are expected to satisfy most of the requirements for full certification.

The left half of Table 3.2 reports the percentage of teachers who lack full certification. In 1990–1991, there is virtually no systematic ordering by income. By 1995–1996, noncertification rates have begun to climb, particularly in schools with a higher percentage of students enrolled in subsidized lunch programs, although some of the increase is likely a result of better reporting of teachers with emergency credentials. There is no doubt, however, that noncertification rates increase dramatically following CSR. In 1997–1998, the noncertification rates were between 4.6 and 5.9 percent for low-poverty schools (0 to 25 percent of students enrolled in subsidized lunch programs), compared with a range of 11.8 to 23.5 percent for high-poverty schools (75 percent or more enrolled in subsidized lunch programs). Similar differences remain in 1999–2000, with even higher percentages of noncertified teachers in all categories. For both 1997–1998 and 1999–2000, nearly all teachers not fully certified are emergency or waiver teachers, whereas only a small portion are interns, compared with a roughly even split between the two groups in 1995–1996.

Within each income category, the increases in noncertification rates from 1995–1996 to 1999–2000 for blacks and Hispanics are noticeably higher than for Asians and whites. In 1999–2000, over 25 percent of black and Hispanic students in high-poverty schools have teachers who lack full certification.

Education

The close link between education and certification described above is reflected in the right half of Table 3.2. Similar to experience and

---

[12]A preintern teacher is a type of emergency teacher.

Table 3.2

Teacher Credentials and Education, by Student Race/Ethnicity and Percentage of Students Enrolled in Free or Reduced-Price Lunch Programs, All Students

	% Not Fully Certified				% with Bachelor's Degree Only			
% of Students in Lunch Program	1990–1991	1995–1996	1997–1998	1999–2000	1990–1991	1995–1996	1997–1998	1999–2000
Asian								
0–25	0.1	0.4	5.5	6.4	12.7	9.0	14.2	13.6
25–50	0.2	0.5	6.5	8.8	14.5	12.9	15.7	16.0
50–75	0.2	1.0	9.7	12.4	19.4	14.8	19.6	20.5
75–100	0.2	2.2	14.3	18.0	25.3	20.4	25.0	29.5
All	0.2	1.1	9.1	10.9	16.9	14.3	18.8	19.4
Black								
0–25	0.3	0.5	5.9	7.7	12.7	11.0	15.4	13.9
25–50	0.7	0.8	7.9	9.9	16.4	14.3	18.8	17.2
50–75	1.2	2.0	12.0	15.5	21.9	15.2	22.2	22.7
75–100	0.4	4.7	23.5	28.3	31.9	27.1	32.5	37.3
All	0.7	2.9	16.3	19.5	22.3	20.0	26.0	27.5
Hispanic								
0–25	0.2	0.3	5.7	6.3	13.3	9.4	13.9	13.6
25–50	0.6	0.8	8.1	8.8	15.5	12.2	16.4	15.5
50–75	0.9	1.7	13.1	14.0	20.1	13.5	21.4	21.2
75–100	0.4	3.5	21.7	25.9	31.5	25.6	32.2	35.1
All	0.5	2.4	16.9	19.6	22.6	19.5	26.5	27.9
White								
0–25	0.1	0.3	4.6	4.5	11.7	9.5	12.4	12.2
25–50	0.4	0.5	5.9	6.3	15.1	11.2	13.6	13.7
50–75	0.4	0.9	8.5	9.4	18.8	12.4	17.0	17.8
75–100	0.3	1.6	11.8	13.8	22.4	17.2	21.9	23.1
All	0.2	0.6	6.5	6.9	14.0	11.3	14.7	14.8
All Students								
	0.4	1.6	12.2	14.3	18.2	15.8	21.3	22.4

NOTE: The percentages are weighted by the number of students in each racial/ethnic category.

certification, the percentage of teachers with only a bachelor's degree rises nearly monotonically with the percentage of students enrolled in subsidized lunch programs. This pattern is evident for the entire decade. The difference between white and nonwhite students grows even larger following CSR, and the percentage of teachers with no postgraduate education is far larger in high-poverty schools for nonwhite students. Over one-third of the teachers for black and Hispanic students, along with nearly 30 percent for Asian students, lack any education beyond the bachelor's degree. The broad-based trend toward fewer teachers with only a bachelor's degree observed between 1990–1991 and 1995–1996 for all demographic categories is nowhere to be seen in 1999–2000.

Teacher Characteristics, by Community Type

The previous tables revealed dramatic changes in experience, certification, and education trends following the implementation of CSR. Not only did average teacher quality as measured by experience, certification, and education decline, but differences by student race/ethnicity and socioeconomic composition tended to widen. An important question is whether urban districts fared worse than the rest of the state, even compared with other districts with similar student demographics. There is a widespread belief that cumbersome bureaucracies, poor facilities, high rates of crime and poverty, relatively low salaries, and many other factors combine to inhibit urban districts from attracting and retaining high-quality teachers. The economic boom concentrated in the large metropolitan areas that raised salaries much more in the private sector than in the public schools and the concomitant increase in housing prices likely exacerbated the teacher hiring difficulties faced by many of these districts.

Table 3.3 reports trends in experience, certification, and education by community type. Because of the smaller sample sizes and the high correlation between race/ethnicity and income within each community type, the table presents results separated only by race/ethnicity. Distributions also separated by income are presented in Appendix Tables B.5 and B.6. Similar to the case for all schools taken as a whole (Tables 3.1 and 3.2), black and Hispanic students in each community type have teachers with less experience, certification, and education than do white and Asian students. Table 3.3 shows that the experience distributions in urban schools are similar to the experience distributions in suburban and rural public schools, with the exception of a larger increase in the proportion of inexperienced teachers for white students.

Table 3.3

Teacher Experience, Credentials, and Education, by Community Type and Student Race/Ethnicity

Student Race/ Ethnicity	% with 0 Years Experience 1990–1991	1995–1996	1997–1998	1999–2000	% with 1 Year Experience 1990–1991	1995–1996	1997–1998	1999–2000
Urban								
Asian	5.5	5.7	11.8	7.9	5.6	5.2	9.9	9.5
Black	7.0	7.8	13.5	9.7	6.0	6.5	10.2	11.7
Hispanic	6.6	7.6	14.4	9.6	6.5	6.1	9.3	11.4
White	5.4	6.5	11.4	7.3	5.3	4.5	8.8	8.1
All	6.2	7.2	13.3	9.0	6.0	5.7	9.4	10.6
Suburban								
Asian	6.0	5.1	11.5	8.0	6.3	5.3	9.2	8.2
Black	8.6	5.9	12.6	9.2	7.4	6.2	9.7	9.1
Hispanic	8.1	6.4	12.9	9.2	7.2	6.1	10.4	9.2
White	5.7	4.7	9.7	6.4	6.4	4.9	9.2	7.1
All	6.6	5.4	11.2	7.8	6.7	5.4	9.6	8.1
Rural								
Asian	7.1	4.4	8.4	5.6	7.0	4.3	10.5	5.9
Black	8.2	4.9	10.8	7.0	8.6	5.2	10.8	7.5
Hispanic	7.5	4.9	11.2	7.3	7.8	5.2	11.0	8.0
White	6.1	4.0	7.2	4.5	6.9	3.7	8.9	5.4
All	6.7	4.4	9.0	5.8	7.2	4.3	9.9	6.6

	% Not Fully Credentialed 1990–1991	1995–1996	1997–1998	1999–2000	% with Bachelor's Degree Only 1990–1991	1995–1996	1997–1998	1999–2000
Urban								
Asian	0.2	1.7	12.1	15.5	18.7	17.4	23.8	24.3
Black	0.4	4.0	21.1	25.5	23.7	22.1	30.3	32.1
Hispanic	0.4	3.4	22.3	26.0	28.0	26.4	34.8	36.2
White	0.2	1.1	9.7	11.1	16.9	17.2	20.9	20.8
All	0.3	2.8	18.4	21.9	23.3	22.8	30.1	31.4
Suburban								
Asian	0.2	0.6	6.9	7.7	15.4	12.4	15.0	16.2
Black	1.1	1.2	8.5	10.1	20.1	17.7	19.3	20.9
Hispanic	0.8	1.6	11.6	13	17.2	13.7	18.1	20.0
White	0.1	0.5	5.8	5.6	13.1	10.2	13.0	13.3
All	0.4	0.9	8.1	8.9	14.9	12.1	15.5	16.7
Rural								
Asian	0.3	0.4	6.5	7.8	13.9	8.9	14.8	12.9
Black	0.7	1.1	9.6	11.4	19.3	11.4	18.3	17.8
Hispanic	0.6	1.1	10.5	13.1	14.7	9.7	17.1	17.2
White	0.4	0.5	4.9	5.5	13.3	8.3	12.1	11.8
All	0.5	0.8	7.4	9.1	13.9	8.9	14.4	14.4

NOTE: The percentages are weighted by the number of students in each racial/ethnic category.

The results in Table 3.3 for certification and education show somewhat more pronounced changes in urban schools than in all public schools (Table 3.2). For all racial and ethnic groups in 1997–1998 and 1999–2000, urban schools have higher percentages of noncertified teachers and teachers with no graduate education than do other schools. For example, 26 percent of urban Hispanic students have teachers not fully certified in comparison with 19.6 percent for all Hispanics as shown in Table 3.2. Appendix Table B.6 suggests that this gap increases for all races/ethnicities as the percentage of students enrolled in subsidized lunch programs rises, showing that it is not driven by student income differences across community types. Postgraduate education rates are also lower in urban districts for schools of a given demographic composition. Thus, it appears that being in an urban district represents an additional burden schools must overcome in attracting and retaining highly qualified teachers.

Trends in the Six Largest School Districts

The difficulties urban schools face in competing for teachers are likely to be even more severe in the largest districts. This section examines changes in teacher characteristics by student demographic group for each of the six largest districts in California as of 1990–1991. Again, because of the smaller sample sizes and the high correlation between poverty and racial/ethnic composition, the tables divide schools only by race/ethnicity.[13]

Table 3.4 shows changes in the percentages of teachers with zero or one year of experience. The share of inexperienced teachers in some of California's largest districts has risen to an alarming level. In 1999–2000, more than 30 percent of the teachers for black and Hispanic students in San Diego were in their first or second year. Almost one-quarter of the teachers for black and Hispanic students in Los Angeles, Long Beach, Oakland, and San Francisco were in their first or second year. In sharp contrast, Asian and white students in these districts—as well as all students in Fresno—have rates of inexperienced teachers quite

[13]The results separated by income are somewhat more dramatic but reveal the same overall patterns as those shown here.

Table 3.4

Teacher Experience in the Six Largest School Districts, by District and Student Race/Ethnicity

Student Race/ Ethnicity	% with 0 Years Experience 1990–1991	1995–1996	1997–1998	1999–2000	% with 1 Year Experience 1990–1991	1995–1996	1997–1998	1999–2000
Oakland								
Asian	5.0	4.6	9.6	6.0	4.4	6.7	10.9	8.5
Black	6.0	5.7	10.8	9.9	3.6	6.2	8.6	10.1
Hispanic	7.9	6.4	11.0	11.6	5.8	6.4	10.5	11.2
White	2.1	3.0	5.1	4.1	2.9	2.6	4.1	5.7
All	5.8	5.4	10.3	9.4	4.0	6.1	9.2	9.9
Fresno								
Asian	8.0	4.0	8.9	5.1	5.8	6.5	11.3	4.5
Black	8.3	5.4	9.2	4.2	5.0	3.8	11.2	3.8
Hispanic	7.9	4.4	9.1	4.6	5.0	5.3	10.5	3.9
White	6.0	3.2	5.0	3.7	3.7	2.7	5.6	2.1
All	7.3	4.1	8.2	4.4	4.8	4.8	9.7	3.6
Long Beach								
Asian	7.3	9.9	13.7	9.4	7.1	8.1	12.4	10.6
Black	8.2	10.0	12.7	10.0	7.2	7.4	13.0	11.4
Hispanic	8.8	10.0	15.2	11.2	7.9	7.7	13.3	11.9
White	5.4	6.1	7.1	6.1	4.1	4.2	9.8	7.4
All	7.5	9.2	13.1	9.9	6.6	7.0	12.5	10.9
Los Angeles								
Asian	4.0	6.1	12.2	8.3	5.2	4.6	8.4	11.3
Black	6.2	7.6	15.5	10.5	6.3	6.6	9.5	13.8
Hispanic	5.5	7.2	15.1	10.0	6.5	6.2	7.4	13.6
White	4.0	6.9	12.6	8.6	5.3	4.6	9.4	11.0
All	5.3	7.1	14.7	9.8	6.2	6.0	7.9	13.3
San Diego								
Asian	9.7	4.1	16.0	9.6	7.2	4.7	5.1	9.9
Black	11.9	6.1	20.4	16.5	8.5	6.8	7.5	14.4
Hispanic	13.0	7.2	21.6	18.4	7.0	8.6	8.0	14.9
White	6.1	3.4	14.3	7.1	5.3	3.4	4.6	8.9
All	9.7	5.4	18.5	13.6	6.6	6.1	6.5	12.4
San Francisco								
Asian	6.6	1.7	1.0	8.0	5.6	4.8	12.8	10.5
Black	8.8	3.1	1.7	9.8	6.4	6.7	14.8	13.7
Hispanic	9.5	2.6	1.6	8.8	7.2	6.5	14.9	12.3
White	5.5	1.4	0.3	6.2	4.8	3.3	9.5	8.3
All	7.5	2.1	1.2	8.3	6.0	5.3	13.3	11.2

NOTE: The percentages are weighted by the number of students in each racial/ethnic category.

comparable to the average California school with a similar demographic composition.

The distribution of the proportion of teachers without full certification presented in the left panel of Table 3.5 differs somewhat from the experience distributions. In some ways, the certification rates for the six largest districts mirror that of the state as a whole: few uncertified teachers in 1990–1991 and 1995–1996, followed by a large increase in 1997–1998 with the class size reduction. Furthermore, by 1999–2000 the percentage of uncertified teachers is noticeably higher for nonwhite students than for white students. However, there is also substantial heterogeneity among districts. In San Diego, only 6.7 percent of the teachers for black students and 9 percent of the teachers for Hispanic students do not have full certification, despite the large numbers of new teachers.[14] In Long Beach, Los Angeles, and Oakland, on the other hand, almost one-third of black and Hispanic students have uncertified teachers.[15] The proportion of teachers without full certification for Asian and white students in Long Beach, Los Angeles, Oakland, and San Francisco also exceed the statewide averages for these two demographic groups.

The final indicator of teacher quality is education, and the right panel of Table 3.5 illustrates a remarkable amount of variation in the proportion of teachers without any postgraduate education. Both Fresno and San Francisco have few teachers without at least some postgraduate education across demographic groups and over time. The rate is also lower in Oakland than in demographically similar schools statewide. In sharp contrast, Los Angeles, Long Beach, and San Diego employ large numbers of teachers lacking postgraduate education, much larger than the statewide averages. In these three cities, as well as in Oakland, nonwhite students were much more likely to have a teacher with only a

---

[14]An alternative possibility is that the certification data are quite poor in San Diego, raising questions about this aspect of the analysis. Similar concerns exist with the 1997–1998 data for San Francisco, where every teacher claims to be fully certified. The analysis in the next chapter of class size, teacher characteristics, and achievement is not affected by any possible problems with the data in San Francisco or San Diego.

[15]Fresno has very few teachers without full certification in any of its schools.

Table 3.5

Teacher Credentials and Education in the Six Largest School Districts, by District and Student Race/Ethnicity

	% Not Fully Certified				% with Bachelor's Degree Only			
Student Race/ Ethnicity	1990–1991	1995–1996	1997–1998	1999–2000	1990–1991	1995–1996	1997–1998	1999–2000
Oakland								
Asian	0.8	3.4	19.5	21.2	5.8	8.5	13.0	10.4
Black	0.4	3.8	23.3	29.2	6.7	8.7	12.5	15.8
Hispanic	0.8	5.5	25.2	32.0	6.5	10.6	13.8	19.8
White	0.1	2.1	8.2	10.2	5.7	4.1	5.8	6.4
All	0.5	4.0	22.2	27.6	6.4	8.7	12.5	15.5
Fresno								
Asian	0	0.4	6.4	4.6	6.3	0	0	0
Black	0	0.5	7.5	4.6	4.5	0	0	0
Hispanic	0	0.4	6.7	4.9	5.8	0	0	0
White	0	0.2	3.5	3.1	3.6	0	0	0
All	0	0.3	6.1	4.5	5.1	0	0	0
Long Beach								
Asian	0.3	10.2	28.9	30.6	18.0	20.4	30.7	30.5
Black	0.4	8.9	28.3	30.9	18.4	18.7	30.5	30.7
Hispanic	0.4	10.7	31.2	33.2	18.8	19.8	32.9	33.7
White	0.3	5.6	16.3	19.3	17.0	13.9	20.0	20.3
All	0.3	9.2	27.7	30.1	18.1	18.5	29.9	30.5
Los Angeles								
Asian	0	2.4	21.9	21.7	38.6	38.6	47.6	47.0
Black	0	4.3	28.8	33.6	39.3	39.1	49.8	51.3
Hispanic	0	3.5	28.8	32.0	41.4	40.9	50.8	51.9
White	0	2.0	20.2	19.2	38.8	38.2	46.7	44.8
All	0	3.4	27.5	30.5	40.6	40.3	50.1	50.9
San Diego								
Asian	0	0.1	1.3	3.6	46.0	51.5	27.2	42.3
Black	0	0	2.3	6.7	49.8	58.4	33.1	50.4
Hispanic	0	0	3.7	9.0	50.6	62.1	34.5	52.5
White	0	0.1	0.8	2.1	42.0	52.3	27.2	40.7
All	0	0.1	2.3	5.9	46.5	56.7	31.0	47.4
San Francisco								
Asian	0	0.1	0	17.3	2.4	1.5	1.3	0.7
Black	0	0	0	22.8	3.2	2.6	2.2	1.1
Hispanic	0	0.2	0	22.4	4.3	2.3	1.9	1.1
White	0	0	0	14.9	2.0	1.5	0.5	1.1
All	0	0.1	0	19.3	2.9	1.9	1.5	0.9

NOTE: The percentages are weighted by the number of students in each racial/ethnic category.

bachelor's degree than were white students. These differences, both between and within cities, persist throughout the decade.

Summary

The tables in this chapter demonstrate both the dramatic changes in teacher characteristics following class size reduction and the importance of race and ethnicity in the distribution of teacher characteristics. Perhaps most important, schools with high percentages of nonwhite and low-income students were much more likely to have inexperienced teachers who lack full certification and postgraduate schooling in 1999–2000 than in 1990–1991, and the racial/ethnic and income gaps clearly widened during the decade. Urban school districts appeared to fare worse than the rest of the state, although there was substantial variation among the six largest districts in the changes in the respective teacher characteristics.

These changes in teacher characteristics are only as important as the link between these characteristics and effectiveness in the classroom. The next chapter provides information on the importance of these shifts in teacher characteristics by investigating the relationship between achievement, class size, and the observable measures of teacher quality. Yet it is important to recognize that teacher quality is determined by myriad factors, of which most are quite difficult to measure. Therefore, the empirical analysis also attempts to gain a better understanding of additional changes in teacher quality not captured by the observable characteristics.

4. CSR and Student Achievement

Introduction

In this chapter, we look at the effect of CSR on student achievement in third grade. In particular, our analysis measures two important effects of CSR on achievement: effects from the reduction in class size and effects from the change in the teacher force. One main finding of the Tennessee STAR experiment was that, all else equal, smaller classes are associated with higher achievement. In California, though, all else was not equal. Chapter 3 describes the dramatic changes in teacher characteristics following the reduction of average class size, particularly for nonwhite students. This increase in the number of inexperienced and uncertified teachers is to be expected when hiring many new teachers, and the negative effect of this increase should disappear as these new teachers acquire experience and certification.

However, CSR likely had a deeper and more pervasive effect on the teacher workforce than simply increasing the number of inexperienced and uncertified teachers. Thousands of additional teaching positions were created, but thousands of additional teachers were not. Thus, much of the increase in the number of teachers consists of individuals who would not have been hired as teachers in the absence of CSR. If these teachers are less capable on average and continue to be less capable even after they have acquired additional experience, certification, or education, then CSR has the potential to create a long-term decline in the average effectiveness of California's teacher force.

Because the relationship between class size, teacher quality, and student achievement is complex, our analysis provides several estimates of the effect of smaller classes. Some estimates combine the effect of smaller classes and new teachers, whereas others provide separate estimates of class size and teacher characteristics. We also investigate whether the relationship between class size and achievement depends on when schools implemented CSR. Once schools have implemented CSR, there is no longer a need to hire additional teachers, and the small year-to-year fluctuations in class size resulting from changes in enrollment are unlikely to be correlated with systematic changes in teacher quality. On the other hand, the pre-CSR/post-CSR difference in class size is accompanied by the hiring of additional teachers. Therefore, estimated effects of smaller classes from these schools are likely to capture any decline in instructional effectiveness that is linked to the magnitude of the reduction in class size needed to satisfy CSR.

Estimation of Class Size Effect

The easiest programs to evaluate are the ones that can be observed as random experiments. In the context of CSR, the ideal experiment (for the purposes of evaluation) would be to randomly assign students to two systems of schools that are identical in every relevant aspect except for the enactment of class size reduction. One set of schools would reduce class size and the other would not. Those reducing class size would have to hire large numbers of additional teachers and the others would not. An important part of the experiment would be that the implementation of CSR could not affect the schools not reducing class size in any way. For example, teachers could not switch from one set of schools to another. Under these conditions, a comparison of the change in academic performance in the two groups of schools over a period of years would provide a valid measure of the total effect of class size reduction: the effect of smaller classes and the effect of hiring new teachers.

To conduct a statewide experiment in a state as large as California is virtually impossible to conceptualize, let alone undertake. In fact, such a comprehensive study of class size reduction has never been attempted. Rather, virtually all research focuses solely on the effects of smaller classes holding everything else, including teacher quality, constant. Although not comprehensive, this research still provides important and policy-relevant information.

The general problem that must be addressed in research on class size is that class size differences tend to be associated with other factors that also affect achievement. Much of the problem comes from the fact that families choose where to live, what schools their children should attend, and in some cases which teacher they prefer in a particular grade. At the same time, principals and other school personnel assign students and teachers to classrooms to achieve a variety of objectives. These choices make it very difficult to separate the causal effect of class size from the confounding influences of other family and school factors.

Consider family decisions first. Families with greater resources and commitment to schooling tend to choose districts and schools with smaller classes or, in this case, those more likely to implement CSR in a timely manner. Thus, the finding that achievement is higher in classes with fewer students or in schools that adopted CSR earlier may be driven primarily by family background. The key point is that these students would tend to do better regardless of class size.

On the other hand, what if policymakers provide additional resources to reduce class size in schools with a high percentage of disadvantaged students? A finding that students in larger classes outperform those in smaller classes would reflect preexisting differences in student preparation rather than perverse effects of class size. Unless appropriate statistical methods are used, it is not possible to identify the contribution of class size or teacher characteristics in a world in which the distribution of students in classrooms and schools is a product of numerous choices by students, parents, teachers, principals, and central administrators. Because the California administrative data and most other datasets contain only limited information on students and teachers, it is highly unlikely that the included variables are able to capture all relevant differences among schools.

Researchers have thus turned to social experiments and innovative statistical methods to identify the causal effect of smaller classes *holding teacher quality constant*. The most prominent example of such an analysis is the Tennessee STAR experiment. Students were randomly assigned to small classes (treatments) or larger classes (controls). A comparison between achievement in large and small classes provides an estimate of the benefits of smaller classes but provides no information on changes in teacher quality. Krueger (1999) and Krueger and Whitmore (2001) find that smaller class sizes in kindergarten and first grade have a significant and lasting effect on achievement. However, Hoxby (2000) discusses potential problems with the randomization of both students and teachers, thereby raising some questions about the validity of the results.[1]

Recent studies by Angrist and Lavy (1999), Hoxby (2000), and Rivkin, Hanushek, and Kain (2000) use quasi-experiments to learn more about the effects of class size on achievement (in Israel, Connecticut, and Texas, respectively). Although these studies do not use true experiments, they are able to isolate changes in class size from other changes in the classroom (and school). Except for Hoxby (2000), the studies find that smaller classes significantly increase achievement in the early grades (grade five and below), and the effects tend to be larger for lower-income students. Because Hoxby's data lead to bias against finding an effect of smaller classes, the overall pattern of results suggests that smaller classes are associated with higher achievement, holding all else, including teacher quality, constant.[2]

Existing Research on CSR

The implementation of CSR in California is a far cry from an ideal experiment. Every school in California was eligible for class size reduction, starting with the 1996–1997 school year. There were large within- and between-school movements of teachers in schools that implemented CSR, as well as in those that did not. Statewide testing did not begin until the spring of the 1997–1998 school year. By this time, most schools had implemented CSR in the first and second grades. However, schools did vary in their timing of CSR in third grade. Nearly one-third of the schools had not implemented CSR in third grade as of 1997–1998. Therefore, it is possible to compare achievement in schools that reduced class sizes in third grade to achievement in those that did not.

---

[1]See also Ehrenberg et al. (2001) for additional discussion of the Tennessee STAR experiment.

[2]In the Connecticut data used by Hoxby (2000), tests are administered in the fall. Therefore, the tests are regressed on class size for the previous school year. If a student moved into the school or district for the current academic year, he or she would cause an erroneous measure of class size. Such measurement error attenuates the estimates of class size effects on achievement.

That is precisely the approach taken by the CSR Research Consortium, which recognized that there may be other differences between early and late adopters of CSR. The consortium 2000 report uses student demographic information and school average fifth-grade scores to control for important differences between early and late CSR adopters that also affect achievement.[3] The use of fifth-grade test scores is logical because fifth-grade achievement captures many of the non–class size differences between schools in achievement. Essentially this method compares the difference between third- and fifth-grade test scores in schools that adopted CSR to the difference in schools that had not yet adopted CSR. A finding that third-grade scores are higher relative to fifth-grade scores in schools that implemented CSR earlier provides evidence in support of the notion that smaller classes raise achievement.

The validity of this method rests on the assumption that the timing of CSR implementation did not affect achievement in fifth grade. The consortium finds substantial increases in the percentage of fourth- and fifth-grade teachers with little or no experience or lacking full certification following the implementation of CSR, compared with much smaller increases for secondary school teachers. Since the total number of fourth- and fifth-grade teachers remained relatively constant, the most logical explanation for this increase is that many fourth- and fifth-grade teachers switched to an earlier grade to take advantage of the smaller classes.[4] The consortium assumes that these changes had no effect on student achievement in fifth grade, in spite of mounting evidence (mentioned above) and the conventional wisdom that novice teachers are associated with lower student achievement.

Instead, the consortium approach attributes any changes in fifth-grade achievement to CSR in third grade. Any decrease in student achievement in fifth grade resulting from a deterioration in instructional effectiveness would increase the gap between third-grade and fifth-grade

---

[3]See Bohrnstedt and Stecher (2000). They use the same approach in their 1999 report, except that they use fourth-grade test scores rather than fifth-grade test scores (see Bohrnstedt and Stecher, 1999).

[4]See Reichardt (2000).

achievement in schools that implemented CSR in third grade, thereby raising the estimated effect of CSR on achievement in third grade. Consequently, this approach cannot distinguish changes in third-grade achievement caused by smaller third-grade classes from changes in fifth-grade achievement that resulted from a decline in fifth-grade teacher quality.

Another limitation of the consortium approach is that it considers achievement in only one school year. For example, the 2000 report uses achievement data only from the 1998–1999 school year. Therefore, this technique is unable to separate the effects of CSR in previous years from other effects such as teacher experience.

The consortium's February 2002 report uses multiple years of data and finds little relationship between length of exposure to CSR and achievement. Unlike previous reports, this consortium report does not use any data at the school or student level; all data are at the state level. The analysis contains a small sample size of fewer than 20 observations, making it very hard to find a systematic relationship between CSR and achievement. The next consortium report, available in the summer of 2002, will use school- and student-level data.

Empirical Approach

Our approach in this report attempts to overcome some of the impediments to the estimation of the effects of CSR and to gain a better understanding of both the pure effects of smaller classes in California public schools and the effects that occurred as a result of changes in the teaching force throughout the state. Rather than comparing the effect of CSR across schools, our analysis identifies class size effects through a comparison of the changes in average achievement and class size over time within schools.[5] We use data for third-grade students in 1997–1998 and 1999–2000. The empirical analysis focuses on test score performance in third grade as measured by the proportion of students (excluding ELs) in a school who exceed the national median test score for

[5]This approach is similar to the one used by Hoxby (2000).

that year.[6] The regression analysis controls for differences in student demographic composition, including percentage enrolled in subsidized lunch programs, percentage black, percentage Hispanic, percentage Asian, and percentage classified as EL. Appendix C provides a more formal discussion of the method used to identify the effects of class size and teacher characteristics on achievement.

This approach ignores achievement differences among schools, thereby avoiding an important source of contamination. It uses only the changes over time in achievement, class size, and teacher characteristics at each school to identify the effects of these variables. For example, if the average gain in achievement between 1997–1998 and 1999–2000 were larger for schools that also experienced larger average reductions in class size, our approach would estimate a positive effect of smaller classes on achievement.[7]

To address the fact that a major class size reduction such as that undertaken by California almost certainly alters the composition and average quality of the teaching force, each model is estimated twice. The first set of estimates comes from a model that does not control for any teacher characteristics. The class size estimates in this model represent the comprehensive (or the "gross") effect of the change in class size, including the effect of class size that we can attribute to teacher attributes such as experience. The second set of estimates comes from a model that accounts for teacher experience, certification, and education. The class size estimates in this model do not contain the effects of teacher attributes, because those characteristics appear explicitly in the model. In other words, these estimates are not comprehensive; they are the "net" effects of class size.

The model providing the "net" estimates of class size also provides estimates of the effect of each teacher characteristic on student achievement. The teacher characteristics include the percentage of novice teachers, the percentage of second-year teachers, the percentage of

---

[6]Individual-level test score data are not available for multiple years. Similarly, the average (at the school level) of individual student test scores is not separately available for non-EL students in 1997–1998.

[7]This method is actually somewhat more complicated in that it controls for differences in other variables included in the regressions.

uncertified teachers, and the percentage of teachers with no postgraduate education. As with class size, we consider only the within-school change in each of these teacher characteristics and its association with the within-school change in student achievement.

Because test score results for individual students are not available, separate school performance measures for each racial/ethnic or income group cannot be calculated.[8] Nevertheless, it is possible to examine whether the effects of class size and teacher characteristics vary according to student demographic composition. Most evidence suggests that class size effects are somewhat larger for lower-income students, although teacher quality changes may counteract any larger effects.[9]

It is also possible to separate the test scores of EL and non-EL students. We do not include test scores for EL students for two reasons. First, because the test is given in English, the test scores are poor indicators of student performance for students with little or no knowledge of English. Students who cannot even read the questions are unlikely to answer them correctly, regardless of academic ability. Second, most EL students were exempted from the test in 1997–1998, but by 1999–2000 all students (EL or not) were expected to take the test to calculate the school's Academic Performance Index. This dramatic change in the composition of EL test-takers makes it very difficult to compare the test scores of any group containing EL students across time.

Our focus on academic achievement as measured by standardized tests does not imply that this is the only outcome of interest. At the elementary school level, these tests are the only measure of achievement that is readily available in California. Plus, they are a more objective

---

[8]Student- or classroom-level data are not available; therefore, achievement, class size, teacher characteristics, and student demographics are measured at the grade level, with the exception of the percentage enrolled in subsidized lunch programs and the percentage classified as EL, both of which are measured at the school level. Although percentage EL at the grade level is available, percentage EL at the school level is preferred. Because principals may systematically assign students to classes partly on the basis of academic performance, classroom-level data would be problematic even if it were available.

[9]See Summers and Wolfe (1977), Krueger (1999), Angrist and Lavy (1999), and Rivkin, Hanushek, and Kain (2000).

measure of achievement across schools and districts than are subjective measures such as teacher evaluations or grades.

One important issue that cannot be resolved is the inability to follow individual students over time. Given the substantial amount of mobility in California, the third-grade cohort in a school contains many students who were not in second grade in that school in the previous year. The result is that the estimated effects of class size and teacher characteristics capture the current year effect plus some portion of the effect from prior years.[10] Therefore, these estimates should not be treated as either the single-year effect of each variable or the sum total of the cumulative effects beginning with kindergarten or first grade.

Findings

The estimated benefits of a ten-student reduction in class size in terms of gains in third-grade mathematics and reading scores are derived from the regression results that are reported in Appendix C.[11] These benefits are presented in a series of figures that depict the estimated change in the percentage of non-EL students who exceed the national median test score in a particular subject following a ten-student reduction in class size.

Following the presentation of the basic results, the analysis investigates the possibility that the benefits of smaller classes differ by racial/ethnic or income composition at the school level. Estimates are computed for the state as a whole, by community type, for Los Angeles, and for the five next largest districts combined.

The final portion of the empirical analysis attempts to gain additional information from the differences in timing of CSR. The

---

[10]In other words, the available California data measure the grade-level achievement for each grade from grades 2 to 11 in each school. However, the data do not measure the substantial student mobility that occurs each year. Therefore, it is not possible to analyze the gain in test scores as students progress through school.

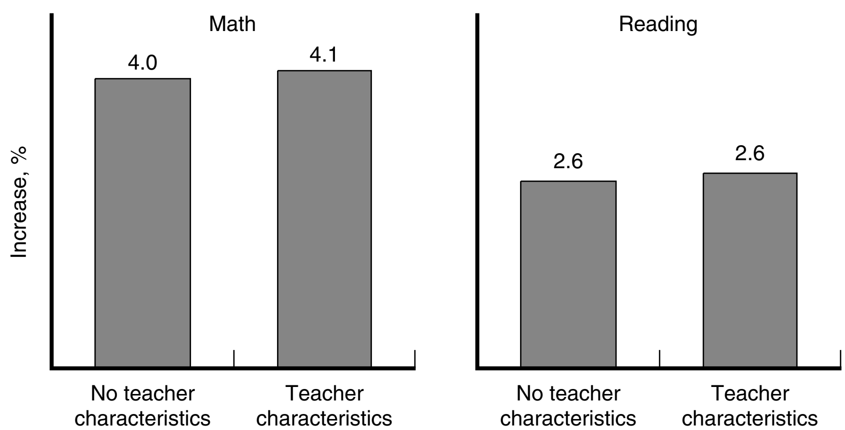
[11]The regression class size coefficients represent the estimated effect of a one-student reduction in class size. Multiplying the coefficients by ten yields estimated gains from a ten-student reduction in class size, which is roughly the average gain across the state. This approach depends upon the assumption that the effects are linear, which is supported by unreported regression results on a large sample of public schools in Texas.

methodology for this examination is described in more detail before the presentation of the results.

Basic Results

Figure 4.1 presents the estimated gains from a ten-student reduction in class size (roughly the average under CSR) on mathematics and reading scores, based on the regression results reported in Appendix Table C.1. The figure shows that a ten-student reduction in class size raises the percentage of students who exceed the national median test score by roughly 4 percentage points in mathematics and less than 3 percentage points in reading, regardless of whether teacher characteristics are included. The 1999 consortium report found similar effects of CSR, but the 2000 consortium report found noticeably smaller effects. The finding that class size effects are larger for achievement in mathematics than in reading mirrors the pattern of findings for fourth- and fifth-grade students in Texas public schools.

The similarity between the class size results including teacher characteristics and the class size results excluding teacher characteristics



NOTE: For bars with dark shading, the effect of class size is significantly different from zero in a statistical sense (using 95 percent confidence intervals).

Figure 4.1—Estimated Increase in the Percentage of Students Exceeding the National Median Test Scores After a Ten-Student Reduction in Class Size

deserves further explanation. One possibility is that these teacher characteristics explain little of the variation in the quality of instruction. This is not the case for teacher experience, as shown by the results in Appendix Table C.1 for the model including teacher characteristics. The estimates indicate that having a novice teacher reduces the percentage of students who exceed the national median by roughly 3 percentage points in both mathematics and reading. However, there is little or no evidence that teacher education or certification is significantly related to the quality of instruction as measured by student achievement. This implies that the dramatic increase in the percentage of teachers without full certification does provide direct evidence of a decline in quality.

A second possibility is that class size differences are not strongly associated with changes in teacher characteristics. Although Chapter 3 documented the dramatic increase in the share of teachers with little or no experience following the implementation of CSR, most schools had already hired additional teachers by 1997–1998. Consequently, changes in experience, education, and certification rates had already occurred in most schools. This was not the case for those that postponed implementation of CSR until 1998–1999 or 1999–2000, and we focus on that subset of schools below to learn more about changes in teacher quality.

Although the finding for teacher education is quite consistent with the bulk of research on the determinants of academic achievement, the existing evidence on certification is much more mixed. For example, Goldhaber and Brewer (2000) find that the students of teachers with subject-matter certification in mathematics perform better than the students of other teachers, and the students of teachers with emergency certification perform no worse than the students of teachers with standard certification. There are concerns with the methodology in most studies concerning teacher certification (as pointed out in Wayne and Youngs, 2001), so the mixed results are to be expected.[12]

To investigate the relationship between achievement and certification in greater detail, we further divide teachers who are not fully certified into two categories: teachers in the process of becoming

---

[12]For example, see Betts, Rueben, and Danenberg (2000).

certified (interns and preinterns) and teachers with emergency or waiver certification.[13] Over two-thirds of the teachers not fully certified in both 1997–1998 and 1999–2000 are emergency or waiver teachers, whereas the remainder are interns. We cannot distinguish between the two types of fully credentialed teachers (preliminary and professional clear), but this limitation is mitigated by controls for the key differences between these two types: experience and education.

The results from a series of alternative specifications for certification and experience are reported in Appendix Table C.2. These models still rely on within-school changes in teacher characteristics and student achievement, but they use different definitions of teacher characteristics. Regardless of the taxonomy used to specify noncertified teachers, there is little or no evidence that certification is systematically related to achievement. Only when teacher experience is excluded from the model does certification become a statistically significant predictor of achievement in both mathematics and reading. In addition, if average years of experience rather than the percentages of teachers in their first and second year is used to measure differences in teacher experience, the magnitude of the certification estimate is larger (and significant for mathematics). The evidence suggests that full certification as presently measured is not a good predictor of student achievement in California elementary schools. However, this conclusion should be qualified because there may be some error in the certification data that could lead to the underestimation of the importance of certification.

Differences by Community Type

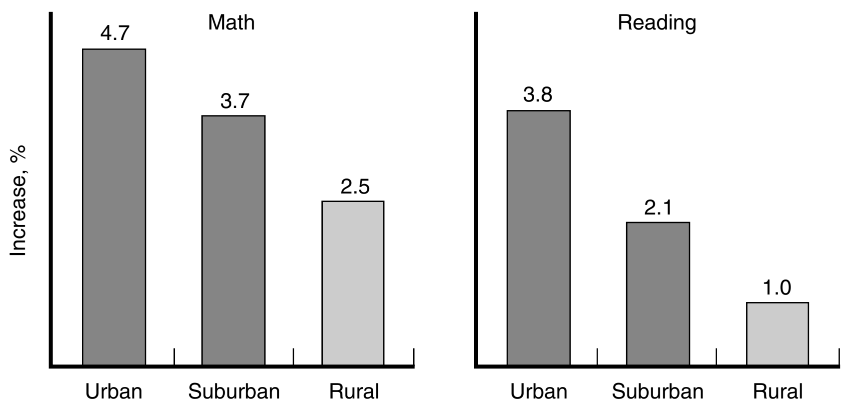
The findings for the state as a whole may conceal substantial differences among schools in the benefits of class size reduction. Appendix Table C.3 reports class size coefficients for urban, suburban, and rural schools as well as for Los Angeles Unified School District and the five next largest districts combined, all derived from separate

---

[13]These two categories are defined as mutually exclusive in our analysis. Teachers who have both emergency/waiver credentials and internship/preinternship credentials are classified only as interns. Similarly, teachers who have full certification and any other form of certification (emergency, waiver, intern, preintern) are classified only as fully certified.

regressions. Figures 4.2 and 4.3 present estimates of the gains from CSR (measured as the effect of a ten-student reduction in average class size) based on the coefficients in the appendix tables.

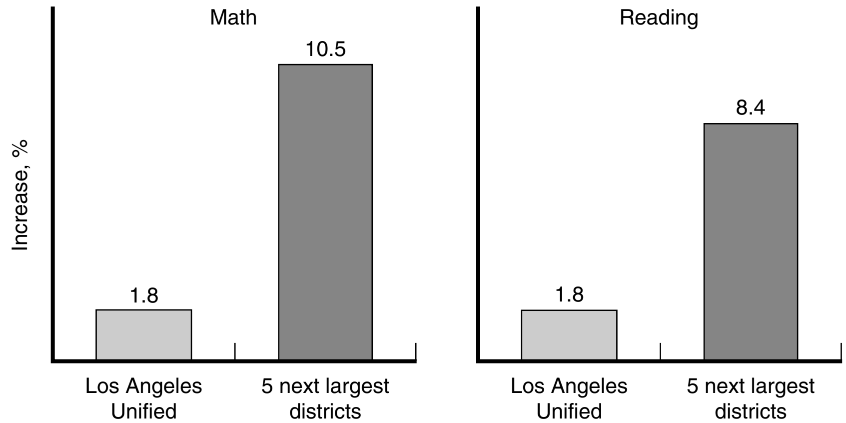
The figures reveal substantial differences in the gains from CSR by community type. Figure 4.2 shows that the average gains in urban schools are 4.7 percentage points in mathematics and 3.8 percentage points in reading; the corresponding gains for suburban schools are 3.7 and 2.1 percentage points in mathematics and reading, respectively. These gains are much larger than the corresponding gains in rural districts, which are not significantly different from zero.[14] Moreover, Figure 4.3 reveals that there is substantial variation in gains even among urban districts. Although CSR led to small and statistically insignificant gains in Los Angeles Unified, the estimated gains in the five next largest districts combined are much larger than the corresponding gains for urban and suburban districts as a whole.



NOTE: For bars with dark shading, the effect of class size is significantly different from zero in a statistical sense (using 95 percent confidence intervals).

Figure 4.2—Estimated Increase in the Percentage of Students Exceeding the National Median Test Scores After a Ten-Student Reduction in Class Size, by Community Type

[14]The reason for the smaller gains in rural schools is not immediately clear. Average class size is quite similar between rural and nonrural schools for 1997 and 1999.



NOTE: For bars with dark shading, the effect of class size is significantly different from zero in a statistical sense (using 95 percent confidence intervals).

Figure 4.3—Estimated Increase in the Percentage of Students Exceeding the National Median Test Scores After a Ten-Student Reduction in Class Size in Los Angeles Unified School District and the Five Next Largest Districts

What factors contribute to the differential gains realized in different districts and community types? One possibility is that the gain from smaller classes depends in part on characteristics of the students. In particular, evidence suggests that lower income children derive larger benefits from smaller classes.[15] If this were the case, differences in student demographic composition among districts would translate into differences in the gain to smaller classes.

However, the fact that the estimated gain is sizable for suburban districts suggests that other factors must also contribute to the differential. Suburban districts on average have a lower percentage of low-income students than do urban or rural districts. The next section of the report investigates differences by student demographic composition for the state as a whole and by community type to learn more about the sources of the differential gains from CSR.

---

[15]See Krueger (1999); Rivkin, Hanushek, and Kain (2000); and Summers and Wolfe (1977). Lazear (2001) provides a theoretical explanation for this result.

Differences by Race/Ethnicity and Income

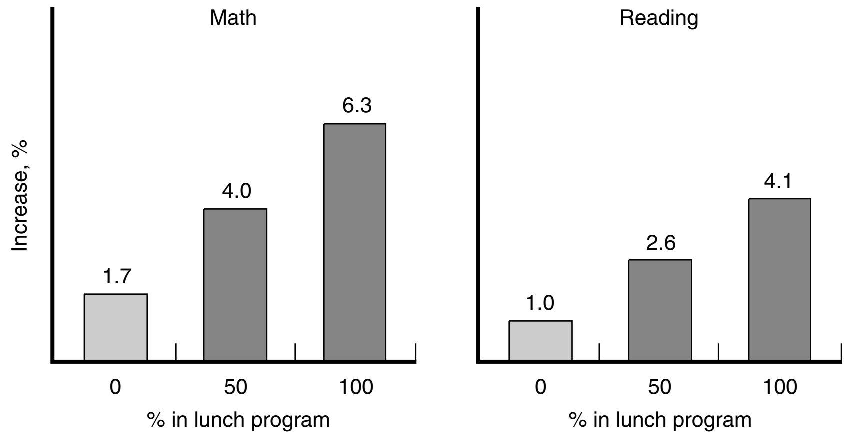
To examine differences by student demographic composition, the regression specifications in this section allow the relationship between class size (as well as any teacher characteristics) and student achievement to differ depending on one of the following two measures of student demographics:[16] percentage of students enrolled in subsidized lunch programs or percentage black.[17] In the previous section, we assumed that the class size effect for schools with no students enrolled in subsidized lunch programs was the same as the class size effect for schools with 100 percent of students enrolled in subsidized lunch programs. In this section, we allow for these two class size effects to differ. The coefficients from these regressions, reported in Appendix Table C.4, are used to construct figures that are similar to those presented above, except that they present three different estimates of the gains from CSR: the gain for schools with 0, 50, or 100 percent of the specified demographic group, respectively. For simplicity, the figures are from the specifications that exclude teacher characteristics. The results in the table show that, as in the previous section, the "gross" and "net" class size effects are quite similar.

Figure 4.4 presents differences in the estimated benefit of smaller classes by the percentage of students enrolled in subsidized lunch programs. Unlike the consortium reports, we find that reducing class size provides an additional benefit for low-income students. CSR raised the percentage exceeding the median mathematics score by slightly less than 2 percentage points in schools that are 0 percent low income but by over 6 percentage points in schools that are 100 percent low income.

---

[16]Because many Hispanics are classified as EL, interactions with percentage Hispanic could potentially provide a very misleading estimate of any differences by the Hispanic enrollment share. EL students are excluded because of both changes in testing criteria over time and the fact that these students tend to receive special instruction for at least a portion of the day outside the classroom for which the included variables are relevant.

[17]Appendix Table C.5 reports results from specifications that combine both interaction effects in the same regression specification. The pattern of estimates is similar to the results from the separate regressions in Appendix Table C.4, but the results in Appendix Table C.5 tend to be much less precise because of the strong association between income and percentage black.



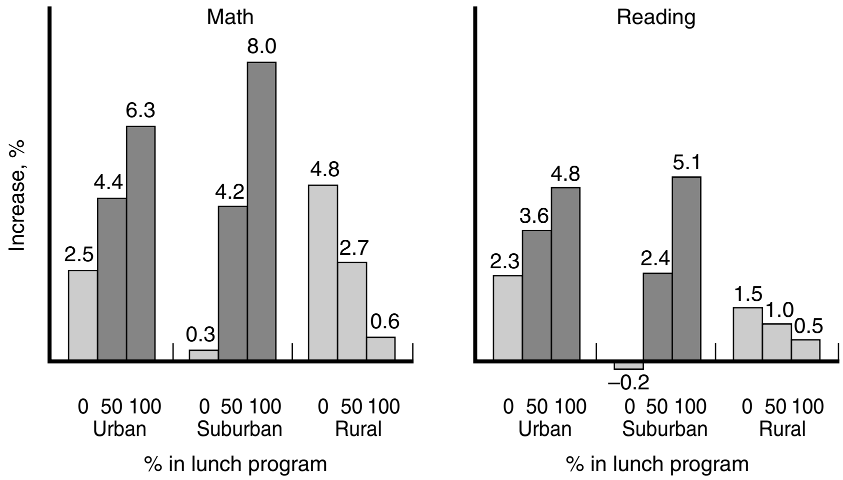
NOTE: For bars with dark shading, the effect of class size is significantly different from zero in a statistical sense (using 95 percent confidence intervals).

Figure 4.4—Estimated Increase in the Percentage of Students Exceeding the National Median Test Scores After a Ten-Student Reduction in Class Size, by Percentage of Students Enrolled in Free or Reduced-Price Lunch Programs

Although the results for reading are not as pronounced, there is a similar pattern.

Figure 4.5 reveals substantial differences in test scores for the percentage of students enrolled in subsidized lunch programs by community type. Among urban and suburban students, lower-income students tend to receive higher benefits from reduced class size. The gains are close to zero in rural districts for both subjects. Figure 4.6 demonstrates that these gains are not found in Los Angeles Unified; in fact, an increase in percentage of students enrolled in subsidized lunch programs is associated with a decrease in the effect of CSR, although the effects are statistically insignificant. In contrast, low-income students in the five next largest districts receive substantial gains from decreasing class size.

Figure 4.7 illustrates differences in the estimated benefit of smaller classes by the percentage of black students in the school. Perhaps most striking is the finding that a higher percentage of blacks actually *decreases* the estimated benefits of CSR on mathematics. The estimated gain is almost 5 percentage points for schools with no black students, only 1



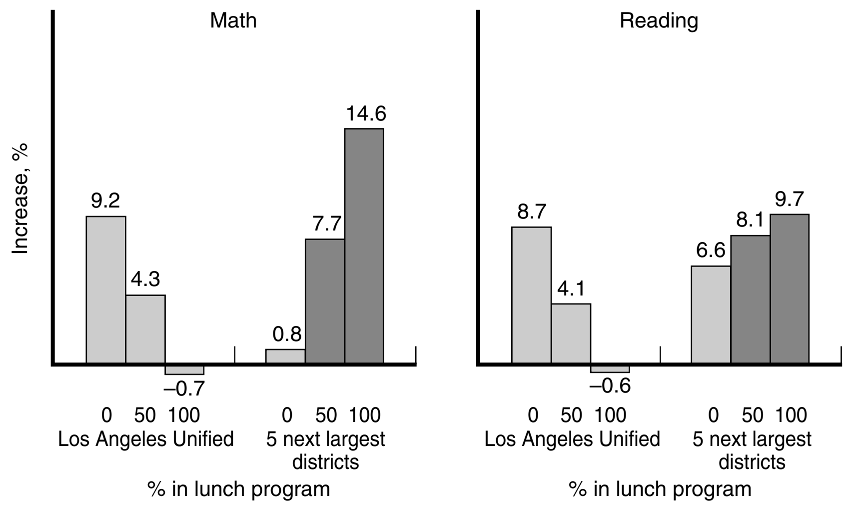
NOTE: For bars with dark shading, the effect of class size is significantly different from zero in a statistical sense (using 95 percent confidence intervals).

Figure 4.5—Estimated Increase in the Percentage of Students Exceeding the National Median Test Scores After a Ten-Student Reduction in Class Size, by Percentage of Students Enrolled in Free or Reduced-Price Lunch Programs and Community Type

percentage point in schools with 50 percent black enrollment, and a negative 2 percentage points (a loss!) in schools with 100 percent black enrollment. The estimated gains from CSR in reading are largely unaffected by the percentage of students who are black.

As shown in Chapter 2, black students predominantly attend urban schools. Figure 4.8 shows that a higher percentage of black students is associated with a noticeably lower gain from CSR in math for urban students, although the estimates are not statistically significant for math or reading. There is little variation in the percentage black for suburban and rural schools, so it is not surprising that there is no discernible difference in class size effects between black and nonblack students in these schools.

Figure 4.9 demonstrates substantial variation even among the large urban districts. Specifically, the negative gain from smaller classes in schools with high black enrollment is concentrated in Los Angeles

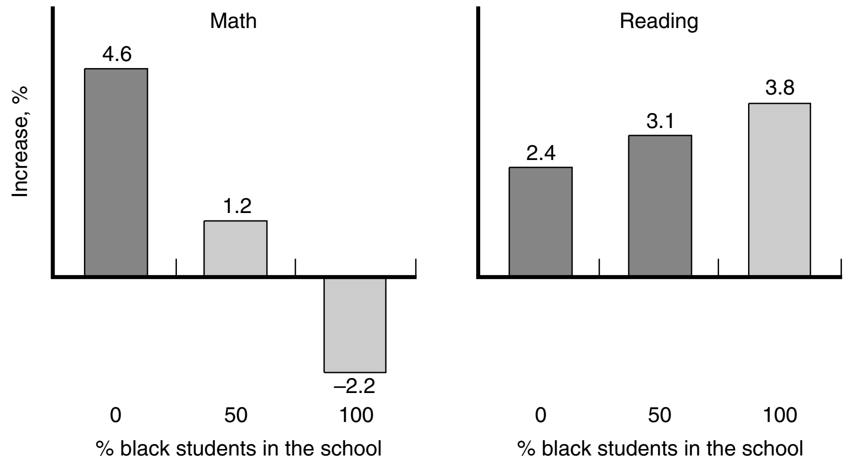


NOTE: For bars with dark shading, the effect of class size is significantly different from zero in a statistical sense (using 95 percent confidence intervals).

Figure 4.6—Estimated Increase in the Percentage of Students Exceeding the National Median Test Scores After a Ten-Student Reduction in Class Size in Los Angeles Unified School District and the Five Next Largest Districts, by Percentage of Students Enrolled in Free or Reduced-Price Lunch Programs and Community Type

Unified School District; the five next largest districts do not exhibit this pattern. In fact, when the sample includes every district except Los Angeles Unified, the negative gain from smaller classes in schools with high black enrollment disappears.

The disparity in findings for blacks is made even more striking by the fact that blacks are much more likely to be classified as economically disadvantaged. These results suggest that the higher benefit of smaller classes that typically accrues to lower-income students is more than offset by some factor related to the percentage of black students in a school. One possibility is that any teacher quality decline is more severe in schools with a high percentage of black students. Whether Los Angeles Unified has more serious problems attracting and retaining teachers in schools with a higher percentage of black students or whether facility



NOTE: For bars with dark shading, the effect of class size is significantly different from zero in a statistical sense (using 95 percent confidence intervals).

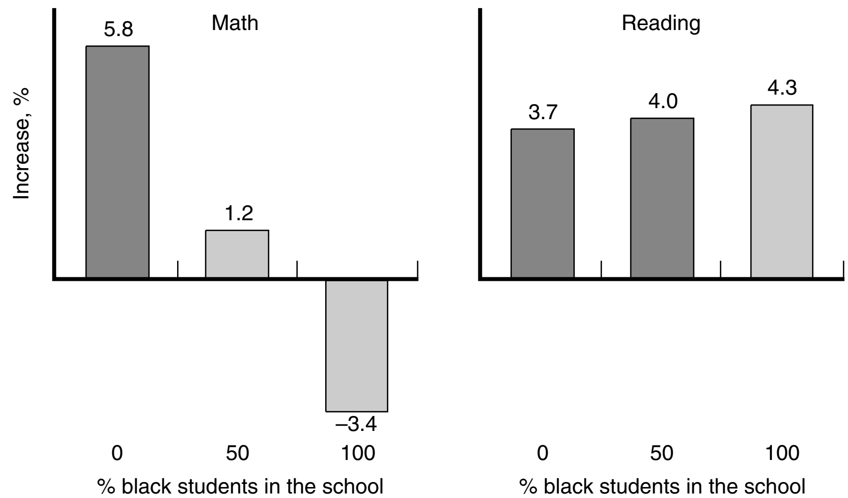
Figure 4.7—Estimated Increase in the Percentage of Students Exceeding the National Median Test Scores After a Ten-Student Reduction in Class Size, by Percentage of Black Students in the School

constraints or other factors inhibit the successful implementation of CSR is an important question to which we now turn.

The Timing of CSR

The results in earlier figures show that most students benefited from CSR, but a substantial number, including those in rural schools, Los Angeles, and schools with a high percentage of black students, gained little if at all from class size reduction. The possibility was raised above that the increase in the number of new teachers may have offset the benefits of smaller classes for these students, and this section investigates this possibility in greater detail.

It is important to note that the analysis to this point provides only partial information on the effect of CSR. Most schools had already reduced class size by 1997–1998, the first year of the sample (and the second year of CSR). However, a subset of schools (over 35 percent) implemented CSR for third grade in 1998 or 1999, and these



NOTE: For bars with dark shading, the effect of class size is significantly different from zero in a statistical sense (using 95 percent confidence intervals).

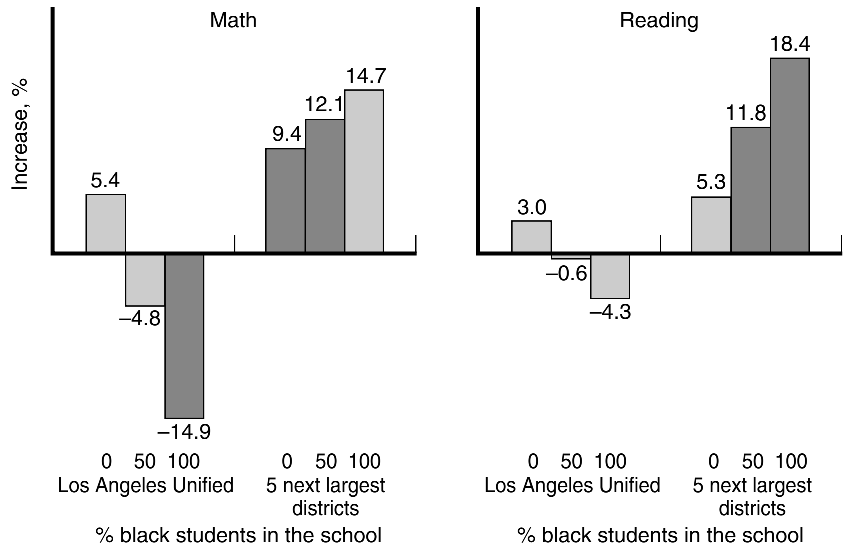
Figure 4.8—Estimated Increase in the Percentage of Students Exceeding the National Median Test Scores After a Ten-Student Reduction in Class Size in Urban Schools, by Percentage of Black Students in the School

schools provide information on the degree to which CSR affected teacher quality.[18] Consider two schools. One (school A) had to reduce average class size by three students to qualify for CSR funds and the other (school B) had to reduce average class size by 12 students to qualify. All else constant, achievement in school B would be expected to increase more than achievement in school A because of the larger decline in class size. However, school B needed to hire proportionately more teachers than school A, and this influx of inexperienced teachers could offset some or even all of the benefits of smaller class size.

As we have noted repeatedly, the challenge in studying CSR is to separate the effect of class size from the effects of hiring more teachers. Our approach is to provide separate estimates of the effect of class size for

---

[18]Because we measure CSR implementation by average class size, rather than from Department of Finance data on participation, our measure of CSR implementation rates differs slightly from that of the CSR Research Consortium (Bohrnstedt and Stecher, 1999 and 2000).



NOTE: For bars with dark shading, the effect of class size is significantly different from zero in a statistical sense (using 95 percent confidence intervals).

Figure 4.9—Estimated Increase in the Percentage of Students Exceeding the National Median Test Scores After a Ten-Student Reduction in Class Size in Los Angeles Unified School District and the Five Next Largest Districts, by Percentage of Black Students in the School

two groups of schools: those that adopted CSR in 1998–1999 or 1999–2000 and those that either adopted CSR before 1998–1999 or had not adopted CSR as of 1999–2000. We also provide separate estimates of the effects of teacher characteristics for the models that include them.

Appendix Table C.6 illustrates the difference in class size between these two sets of schools. For the early adopters (and nonadopters), average change in class size for these schools approaches zero and the standard deviation is also small. This is not to say that there was no variation in class size within this group. Schools receive no CSR money for classes of 21 or more, but schools cannot precisely predict what their enrollment will be for the entire school year. Therefore, schools will often have class sizes of 18 or 19 rather than 20.

On the other hand, Appendix Table C.6 also shows that the average change in class size for the late adopters of CSR (1998–1999 or 1999–2000) was more than eight students. Consequently these schools had to add a large number of third-grade teachers to staff the additional classrooms. These differences in class size and teacher hiring suggest that the estimated effect of class size and teacher characteristics will also be different for these two sets of schools. Both sets of schools experience year-to-year fluctuations in enrollment and teachers, and these changes will be reflected in the class size estimates for both sets of schools. But the schools adopting CSR in 1998–1999 or 1999–2000 have the additional effects on average class size of the implementation of CSR.

Furthermore, there is reason to believe that these schools may have had a harder time hiring extra teachers than schools implementing CSR previously. As mentioned above, the increase in teaching positions was not accompanied by much of an increase in the number of qualified teachers. Therefore, the schools hiring new teachers immediately probably had more teachers from whom to choose. Before CSR, there were fewer than 4,000 new K–3 teachers each year in California; in 1997–1998 that number was over 10,000.

As in the previous section, both the "gross" and "net" estimates of the class size effect are obtained. The "gross" estimates are the comprehensive effects of the change in class size, including the effect of class size that we can attribute to teacher attributes such as experience. The "net" estimates do not include the effect of class size that we can attribute to teacher experience, certification, and education. These teacher characteristics appear elsewhere in the model.

Although we provide separate estimates of class size and (for the "net" estimates) teacher characteristics, this is not equivalent to saying that the late adopting schools are as good on average as the early adopters. As in previous sections, the class size estimates are not identified from achievement and class size differences between the two sets of schools. Instead, the analysis focuses solely on the relationship in each school between the change in average class size and the change in achievement. Any systematic differences in facilities, finances, student backgrounds, or administrator quality are accounted for,[19] and only other factors that are correlated with the change in class size could influence the class size estimates.

Appendix Table C.7 reports third-grade mathematics and reading class size coefficients for the two sets of schools. For each test, the left column contains the "gross" estimated effects of class size from specifications that exclude teacher characteristics, and the right column contains the "net" estimate effects of class size from specifications that include teacher characteristics.[20] The table presents statewide coefficients (i.e., no interaction terms) as well as results that allow the class size coefficient to vary with the percentage enrolled in subsidized lunch programs or percentage black. Only estimates for the entire state are reported, because there are too few observations for late adopters within each community type or district. The figures in the text of this section are constructed from Appendix Table C.7 and, except where noted otherwise, come from the results that exclude teacher characteristics.

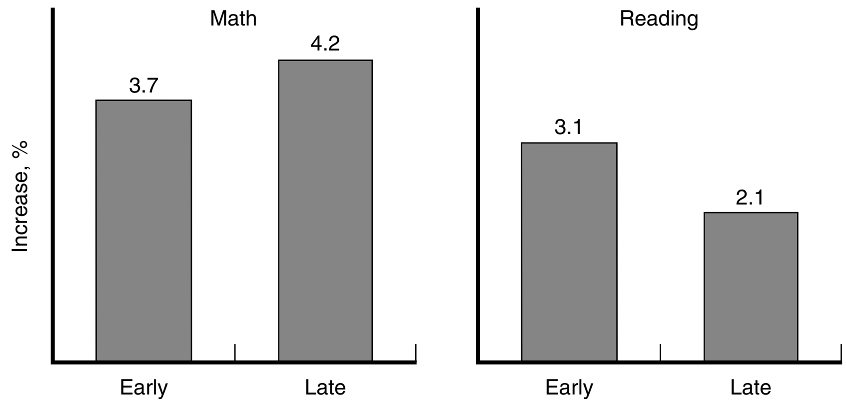
Figure 4.10 provides mixed evidence on whether the hiring of new teachers as a result of the implementation of CSR reduces the overall benefits of smaller classes. In the case of mathematics, the class size effect is slightly larger for the schools implementing CSR in 1998–1999 or 1999–2000, whereas in the case of reading, the class size effect for these schools is roughly 30 percent smaller.[21] Figure 4.11 suggests that the

---

[19]For example, late CSR adopters have higher percentages of students enrolled in subsidized lunch programs (52 percent) and Hispanic students (40 percent) than other schools (45 percent enrolled in subsidized lunch programs and 33 percent Hispanic), although average percentage black is quite similar (roughly 9.5 percent) in schools that implemented CSR in 1998–1999 or 1999–2000 and the others.

[20]One concern with the validity of the results in the table is that the group of schools that did not implement CSR between 1997 and 1999 is a diverse group. It contains schools that implemented CSR before 1997 and those that implemented CSR after 1999. Although not reported, similar results to those in the table are found when we separate the sample into three sets of schools based on the timing of CSR implementation: before 1997, between 1997 and 1999, and after 1999. We do find stronger class size effects for the schools that implemented CSR after 1999 than for schools implementing before 1997.

[21]It should be noted that the differences between early and late adopters are not statistically significant. The quite small variation in class sizes for the sample that did not institute CSR between 1997 and 1999 almost certainly contributes to the rather larger standard errors for the estimated variable effects for this group of schools.



NOTE: For bars with dark shading, the effect of class size is significantly different from zero in a statistical sense (using 95 percent confidence intervals).

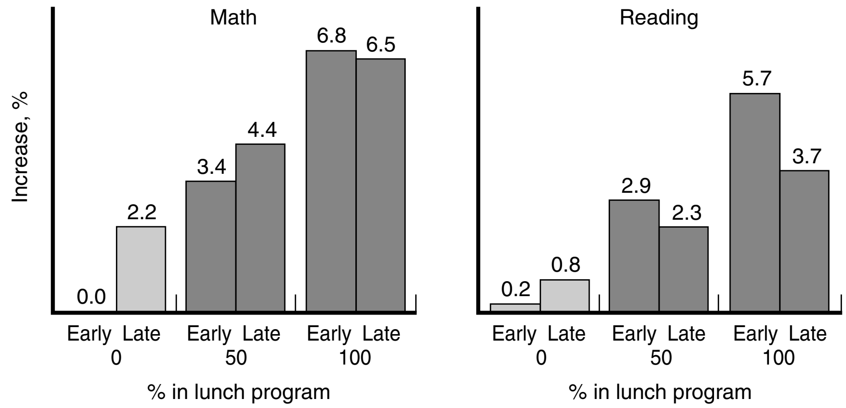
Figure 4.10—Estimated Increase in the Percentage of Students Exceeding the National Median Test Scores After a Ten-Student Reduction in Class Size, by Timing of CSR

results for reading are driven by lower-income students. In the extreme case where all students are enrolled in subsidized lunch programs, the class size effect is much smaller for the schools that implemented CSR in 1998–1999 or 1999–2000 than for other schools.[22] The gap is much smaller but still exists for schools where half the students are enrolled in subsidized lunch programs.

Figure 4.12 presents class size effects from specifications that allow for different class size effects depending on the school's percentage of black students. Because the specifications in the figure exclude teacher characteristics, the effects of any accompanying changes in experience, certification, or education are captured by the class size coefficients. Although evidence based on all students or even allowing for income differences is mixed, the results in Figure 4.12 provide strong support for the belief that gains from implementing CSR fell as the percentage of

---

[22]For mathematics, the coefficients for class size interactions with percentage enrolled in subsidized lunch programs are 20 to 35 percent smaller for the sample of schools that implemented CSR in 1998–1999 or 1999–2000.

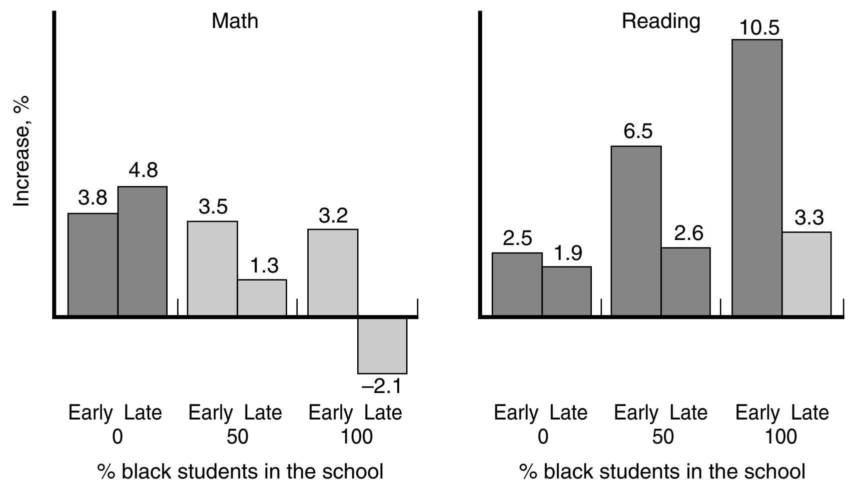


NOTE: For bars with dark shading, the effect of class size is significantly different from zero in a statistical sense (using 95 percent confidence intervals).

Figure 4.11—Estimated Increase in the Percentage of Students Exceeding the National Median Test Scores After a Ten-Student Reduction in Class Size, by Timing of CSR and Percentage of Students Enrolled in Free or Reduced-Price Lunch Programs

black students in the school increased. In those schools that did not implement CSR in 1998–1999 or 1999–2000, the estimated increases in the percentage who exceeded the national median mathematics score are 3.8, 3.5, and 3.2 percentage points for schools that are 0 percent, 50 percent, and 100 percent black, respectively. In contrast, the corresponding numbers for those schools that implemented CSR in 1998–1999 or 1999–2000 are 4.8, 1.3, and –2.1 percentage points. These schools had to increase substantially the number of third-grade teachers. Essentially, the results indicate that students in schools that were at least 70 percent black derived no improvement on average in mathematics achievement following the implementation of CSR, and those in virtually all black schools may even have been hurt.

The estimates do not indicate that students in higher percentage black schools received no benefits from smaller classes. For schools not implementing CSR in 1998–1999 or 1999–2000, the class size estimates



NOTE: For bars with dark shading, the effect of class size is significantly different from zero in a statistical sense (using 95 percent confidence intervals).

Figure 4.12—Estimated Increase in the Percentage of Students Exceeding the National Median Test Scores After a Ten-Student Reduction in Class Size, by Timing of CSR and Percentage of Black Students in the School, Excluding Teacher Characteristics

for mathematics show that blacks benefited as much from smaller classes as other students. However, the average benefit of smaller classes for students in high-percentage black schools that implemented CSR in 1998 or 1999 is quite small or even negative, consistent with the notion that an increase in the number of teachers led to a decline in instructional effectiveness.[23]

Although not as dramatic, the pattern of estimates in Figure 4.12 for reading is similar to that in mathematics. In schools that did not implement CSR in 1998–1999 or 1999–2000, the estimated increases in the percentage who exceeded the national median reading score are 2.5, 6.5, and 10.5 percentage points for schools that are 0 percent, 50

[23]This gap in class size effects based on the timing of CSR implementation also exists when we exclude Los Angeles Unified from the sample, although the negative effect for 100 percent black schools disappears.

percent, and 100 percent black, respectively, whereas in schools that implemented CSR in 1998 or 1999, the corresponding numbers are 1.9, 2.6, and 3.2 percentage points.

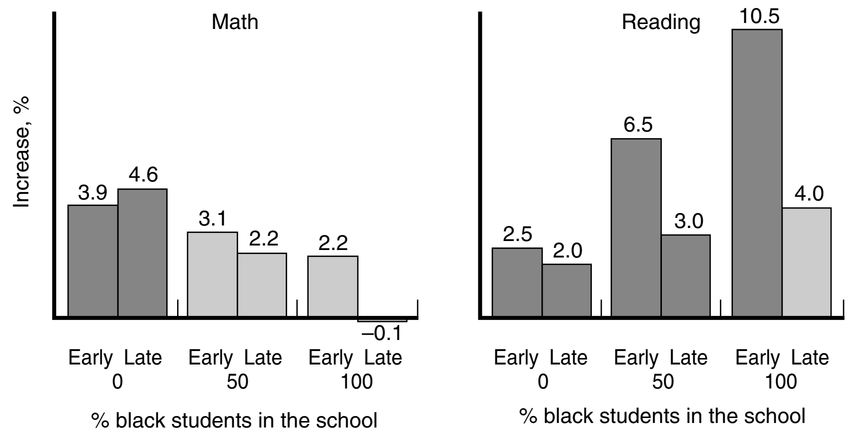
Given the difficulties faced by new teachers, teacher quality would be expected to fall in schools just implementing CSR even if they continued to hire from the same pool of applicants because of the need to hire large numbers of new teachers. This is an unavoidable short-term cost of reducing class size that would disappear as these new teachers acquired more experience. If the apparent decline in the benefits of smaller classes were to result entirely from the lack of experience of the recent hires, this would not be considered a serious problem.

The estimates in Appendix Table C.7 do in fact show that experience accounts for a portion of decline in the benefits of smaller classes. Controlling for experience raises the estimated gains of CSR for students in schools that adopted the program in 1998–1999 or 1999–2000 (the other characteristics do not have a significant effect on achievement).

Whereas Figure 4.12 contains the "gross" effect of class size for schools with different percentages of black students, Figure 4.13 presents the "net" class size effect (i.e., the effect accounting for teacher characteristics). Consider the schools with at least 50 percent black students. If all the decline in achievement between early and late adopters in Figure 4.12 were due to teacher experience, education, and certification, then there would be no decline in Figure 4.13. On the other hand, if none of the decline in Figure 4.12 were due to teacher characteristics, then the gap between early and late adopters in Figure 4.13 would match that in Figure 4.12. A comparison of the two figures suggests that a noticeable portion of the decline in class size effects is attributable to teacher characteristics (largely experience), but more than half the gap still remains.[24]

---

[24]In mathematics, the gap appears to be driven by Los Angeles Unified, because the gap disappears when we exclude Los Angeles Unified from the sample. The gap for reading is still sizable when we exclude Los Angeles Unified but is smaller than in the overall sample.



NOTE: For bars with dark shading, the effect of class size is significantly different from zero in a statistical sense (using 95 percent confidence intervals).

Figure 4.13—Estimated Increase in the Percentage of Students Exceeding the National Median Test Scores After a Ten-Student Reduction in Class Size, by Timing of CSR and Percentage of Black Students in the School, Including Teacher Characteristics

What other factors could explain this gap in class size benefits? As mentioned above, such factors would have to be correlated with the within-school change in class size. The model accounts for CSR implementation, as well as for differences between schools. CSR resulted in the hiring of many teachers, and there is concern that these teachers were less capable than other teachers. The fact that observed measures of teacher quality (experience, certification, and education) account for some but nowhere near all the decline in the benefits of smaller classes suggests that the implementation of CSR reduced the average quality of new hires from what it had been in the past, especially in schools with a high percentage of black students.

Because the variation in class size is much smaller for schools not implementing CSR in 1998–1999 or 1999–2000 (see Appendix Table C.6), the coefficient on the class size/percentage black interaction term is not precisely estimated. Therefore, we cannot reject, even at the 10 percent level, the hypotheses that the mathematics and reading class size effects for high-percentage black schools that adopted CSR in a timely manner equal the effects for late adopters. Nevertheless, the significantly positive interaction term for late adopters combined with the very small interaction coefficient for other schools and prior evidence that blacks benefit as least as much from smaller classes as other students provide support for the view that the benefits of CSR were much smaller for blacks in late-adopting schools.[25]

The Effects of CSR on Fifth-Grade Achievement

CSR implementation in third grade also affects achievement in other grades. The consortium documents large changes in teacher experience in fourth and fifth grades following the implementation of CSR. These changes suggest that many teachers moved to lower grades to take advantage of the now smaller classes. In this case, at least a portion of any decline in teacher quality resulting from CSR would occur in fourth and fifth grades despite the fact that these grades did not reduce average class size. Therefore, a comparison of third-grade class size coefficients might actually underestimate the schoolwide decline in teacher quality from CSR.

To provide additional evidence of the relationship between teacher quality and class size effects, we consider the relationship between changes in third-grade class size and fifth-grade achievement. As in the previous section, the models in this section allow for the effect of class size to differ depending on the percentage of black students in the school. However, the model estimated in this section differs from the one estimated in the previous section in two ways. First, all the variables that were previously estimated at the third-grade level (such as achievement) are now estimated at the fifth-grade level. Second, the model includes the change in third-grade class size in addition to the change in fifth-grade class size. For schools that adopted CSR in 1998–1999 or 1999–

---

[25]Appendix Table C.8 shows that the inclusion of the subsidized lunch interactions increases the magnitude of the coefficient on the percentage black interaction for students in late-adopting schools. This is consistent with the notion that a decline in teacher quality offsets the higher benefits of smaller classes accruing to economically disadvantaged students.

2000, the change in third-grade class size is correlated with how many new teachers the school added to implement CSR.

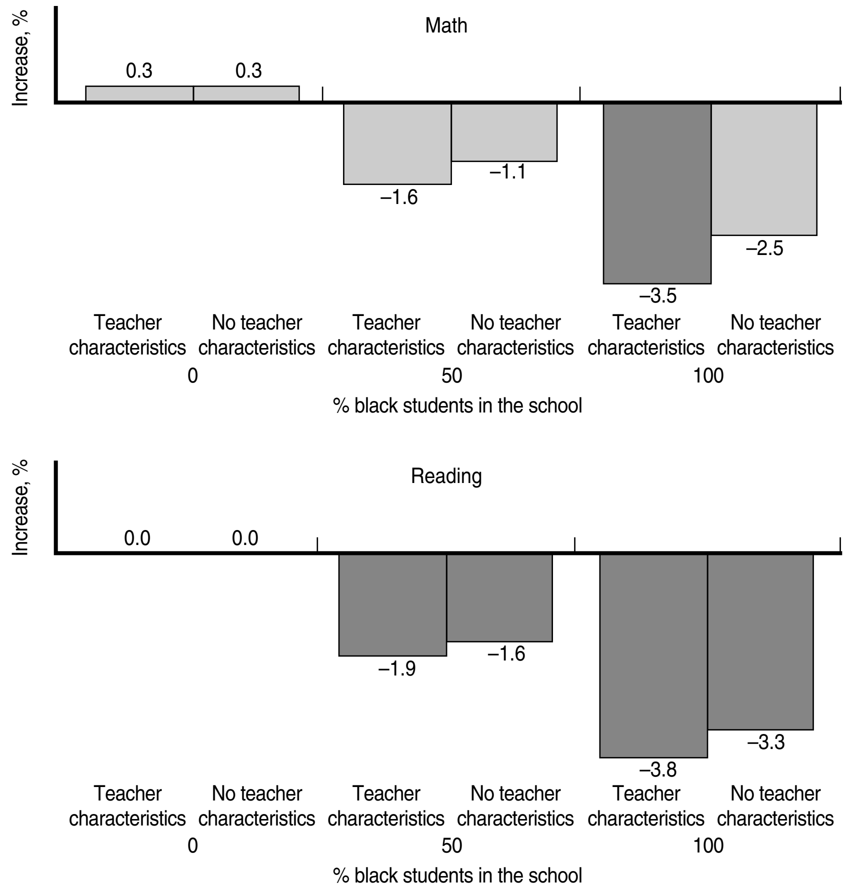
Appendix Table C.9 contains the results from these regressions for fifth-grade achievement that include third-grade class size as well as fifth-grade class size and other controls.[26] The results display a very similar pattern to the results for third-grade test scores. Figure 4.14 shows that a ten-student reduction in third-grade class size is associated with roughly 1 to 2 percentage point drops in the percentages of fifth-grade students who exceeded the national medians in mathematics and reading in schools with 50 percent black enrollment. The declines for schools with 100 percent black enrollment are between 2 and 4 percentage points. In addition, it appears that changes in the percentages of fifth-grade teachers with zero and one year of experience account for roughly one-quarter of the decline associated with the reduction in third-grade class size, as the inclusion of teacher characteristics reduces the negative effect of third-grade class size by roughly one-quarter in Appendix Table C.9. Similar to the evidence for third grade, the findings for fifth grade suggest that in higher-percentage black schools, CSR led to both the need to hire a number of inexperienced teachers and a decline in the average quality of new hires.[27]

The conclusion that a decline in teacher quality offsets the benefits of smaller classes in high-percentage black schools is not based on observable measures of quality such as experience or certification. Rather, it is based on the pattern of the relationship between achievement, class size, and teacher experience. If teacher experience can explain a sizable portion of the decline in the benefits of smaller class size, it seems reasonable that unobserved measures of teacher quality can also explain a portion of that decline. We believe that an overall decline in teacher quality provides the most plausible explanation for the smaller benefits of class size reduction for late adopters of CSR that are only

---

[26]Appendix Table C.10 presents the descriptive statistics for the fifth-grade sample, by CSR timing.

[27]The interaction between percentage of black students in a school and class size is nearly identical when we remove Los Angeles Unified from the sample. Thus, although the negative effects of class size for blacks appear to be concentrated in Los Angeles Unified, the effects for fifth grade appear to occur throughout the state.



NOTE: For bars with dark shading, the effect of class size is significantly different from zero in a statistical sense (using 95 percent confidence intervals).

Figure 4.14—Estimated Increase in the Percentage of Students in the Fifth Grade Exceeding the National Median Test Scores After a Ten-Student Reduction in Class Size in Third Grade, by Percentage of Black Students in the School

partially explained by the influx of new teachers and the link between third-grade class size and fifth-grade achievement.

There is substantial reason to believe that our analysis actually underestimates the decline in average teacher quality both for the state as a whole and for students in those schools that tend to have the most difficulty attracting and retaining teachers. Teacher quality may also be affected by hiring at neighboring schools or other schools in the same district. Consider two schools in Los Angeles. One is in a middle-class neighborhood in the San Fernando Valley and the other is in a high-poverty area of South Central Los Angeles. Class size reduction in the San Fernando Valley school opens up additional jobs that may be filled by teachers currently working in South Central if they have the appropriate level of seniority. There may not be any decline in quality in the San Fernando Valley school, but teacher quality could fall at the school in South Central if that school is unable to find suitable replacements. Because any such decline in teacher quality is not linked to the magnitude of class size change at the affected school, it will not be captured by the methodology used in this report. In fact, no methodology could capture this effect with the currently available data on teacher transitions. Our analysis of third-grade class size effects will understate any decline in teacher quality associated with the program because it does not measure the decline between schools or in other grades.

Summary

A ten-student reduction in class size (the average under CSR) raises the percentage of third-grade students who exceed the national median test score by roughly 4 percentage points in mathematics and 3 percentage points in reading. Schools with more low-income students likely receive larger benefits, whereas schools in rural areas and those in which a high proportion of the students are black (primarily in Los Angeles Unified School District) appear to benefit little if at all from smaller classes.

The relationship between teacher characteristics and achievement is much weaker. The only indicator that is systematically linked to student achievement in third grade is experience. Having a new teacher reduces the percentage of students who exceed the national median by roughly 3 percentage points in both mathematics and reading. There is little or no evidence that teacher education or certification is significantly related to student achievement in third grade. However, the finding for certification could be influenced by the lower quality of the certification data.

One possible explanation for the variation in class size effects is the timing of CSR. For schools with a large percentage of black students, schools recently implementing CSR have much smaller benefits of class size reduction than do other high-percentage black schools. It is likely that the smaller effects of class size for recent CSR implementers are related to changes in hard-to-measure teacher attributes, given that they are related to easier-to-measure teacher attributes such as experience. The schools that do not appear to benefit from CSR are the same schools that had trouble hiring experienced, certified teachers before CSR. In contrast, there is little evidence of a difference in class size benefits based on CSR timing in schools serving predominantly middle class, nonblack students.

5. Summary and Policy Implications

This report examines the effect of class size reduction on third-grade mathematics and reading achievement in California. The now-standard class size of 20 students for grades K–3 is roughly one-third smaller than the average just six years ago. The results suggest that, all else equal, most students benefit from smaller classes and that lower-income students likely receive larger benefits (except in Los Angeles). However, there is substantial variation across districts, and students in rural areas and those in Los Angeles Unified School District appear to benefit little if at all from the smaller classes. In Los Angeles Unified, the struggles to benefit from CSR are particularly challenging for schools with high percentages of black students or students enrolled in subsidized lunch programs.

Although the analysis does not provide a definitive explanation for the failure of CSR to help some of the state's most disadvantaged students, the pattern of findings strongly suggests that a decline in the average effectiveness of California's teaching force is an important factor. Chapter 3 shows that the rapid expansion of the teaching force needed to staff the additional classes led to a dramatic increase in the percentage of teachers who lacked full certification, who had no postgraduate education, and who were in their first or second year of teaching. These teachers were concentrated in schools with high percentages of nonwhite students enrolled in subsidized lunch programs. For example, in 1999, nearly 30 percent of the teachers in these schools were not fully certified, and over 10 percent were novice teachers. In Los Angeles alone, over 30 percent of the teachers lacked full certification and almost one in four were in their first or second year of teaching. Overall, the income and racial/ethnic gaps in teacher quality widened along virtually all measured characteristics between 1995 and 1999.

The increase in the number of teachers with little or no experience is a large but unavoidable cost of class size reduction on the huge scale that California implemented it. Although the lower certification rate would appear to signal a decline in the quality of new teachers, the empirical analysis does not reveal a strong relationship between achievement and teacher certification. However, the mere fact that many schools and districts found it so difficult to hire and retain certified teachers before CSR strongly suggests that the overall quality of the new teachers hired by these districts because of CSR is lower than in other districts.

An example of this difference in teacher quality is found for schools with a high percentage of black students. Our regression analysis of these schools shows that these students benefited from smaller classes, but much if not all of these benefits were offset by what appears to be a decline in teacher quality that accompanied the implementation of CSR. Moreover, the overall decline in teacher quality following CSR is only partially measured by our analysis of third-grade achievement. The analysis of fifth-grade achievement shows that the implementation of CSR in third grade has a negative effect on achievement in fifth grade. The likely source of this decline is the movement of many fourth- and fifth-grade teachers into the early grades.

Our findings in this study have numerous implications for policymakers. The inability of CSR to raise achievement for some of the state's neediest students suggests that the statewide implementation of CSR was far from ideal. It is not surprising that the schools that had difficulty hiring before CSR had even more difficulty hiring after CSR created thousands of openings at schools with fewer needy students. If, instead, these struggling schools had been allowed to reduce class size first, they might have had an easier time attracting and retaining teachers because of their smaller class sizes (compared with other schools). Furthermore, our finding that schools with few low-income students received few benefits from smaller classes implies that delaying CSR in these schools would not have hurt these students.

The most obvious consequence of CSR was the need to hire 25,000 additional teachers, and there is concern that these teachers were less capable than other teachers. For schools with a sizable percentage of black students, the "gross" class size effect (i.e., the effect of smaller classes and of teacher effectiveness) is lower in schools recently adopting CSR. Observed measures of teacher quality (experience, certification, and education) account for some but nowhere near all this decline in the benefits of smaller classes. The most plausible explanation for the remainder of the gap is a change in unobserved teacher quality. Whatever the reason, this gap suggests that the adverse effects of CSR implementation may not simply go away when these new teachers acquire more experience.

The dramatic increase in new teachers will drive up the cost of CSR. As these teachers gain experience, they also receive more pay. Currently, the state funding for CSR does not cover the increase in teacher salaries. The political popularity of smaller classes may make it quite difficult to reduce expenditures through small increases in class size, possibly forcing administrators to take resources from other uses. The Irvine school district has chosen to drop CSR rather than use its own money, and other districts may follow its example.

Our results concerning new teachers highlight the importance of policies targeting new teachers, such as the BTSA. These types of programs have two potential benefits. First, they can minimize the adverse effects of new teachers by helping them adapt to the classroom more quickly and effectively. Making these teachers more productive and effective also reduces the stress of the job, thereby reducing teacher turnover. Effective programs that assist new teachers will become even more essential in the near future, as enrollment and teacher retirements both increase. Future research should carefully analyze the effectiveness of programs such as BTSA on student achievement.

Although we find that experience matters, the relationship between certification and achievement is much less clear. Our results show that California's certification system in the late 1990s has little if any relationship to student achievement, suggesting that policies that prevent uncertified teachers from teaching are unlikely to raise student achievement. Indeed, such policies could have the perverse effect of lowering achievement by preventing talented but uncertified teachers from teaching. For example, Raymond, Fletcher, and Luque (2001) find that test scores in Houston are at least as high, if not higher, for uncertified Teach for America teachers than for new, certified teachers.

The concern about the weak relationship between certification and student achievement is well known to California policymakers, and the California Commission on Teacher Credentialing (CCTC) is currently reforming the certification process. The first goal of these reforms is to align the certification requirements with the content in the classroom to increase the association between certification and achievement. The second goal is to streamline the certification process to increase the number of certified teachers. It is hoped that the effect of these reforms will be studied rigorously.

At the same time, the available data in California on certification limit the inferences that can be drawn. For example, it is not possible to distinguish between the two types of full credentials (preliminary versus professional clear). Nor can the data from the CCTC be linked to the California Basic Education Data System (CBEDS) from the Department of Education. It is not possible to match individual student test scores from one year to the next to measure student growth. Thus, several improvements to the current data collection system are needed to conduct an in-depth evaluation of the effects of certification on student achievement.

Was CSR a good investment? The way it was implemented makes the question extremely difficult to answer. In fact, the effects of CSR will probably never be known with any certainty, particularly in kindergarten, first, and second grades. Our analysis focuses on achievement in third grade. Unlike previous work, it considers the effect of hiring many new teachers in addition to the effects of smaller classes. However, the data on teacher transitions within and between schools are limited, so our analysis probably underestimates the decline in instructional effectiveness. Overall, the gain in the percentage of students scoring above the national median associated with smaller classes is about equal to the loss in test scores associated with new teachers. Schools with many low-income students appear to benefit more from smaller classes, but rural schools, schools in Los Angeles, and schools with high percentages of black enrollment do not gain from smaller classes. Their inability to gain is only partially explained by the increase in new teachers resulting from CSR. Only time will tell if these schools *ever* gain from CSR.

Appendix A

Data Sources[1]

The majority of the data for this study are from the California Department of Education. This appendix provides a description of these data.

CBEDS

CBEDS is maintained and supported by the Educational Demographics Unit in the California Department of Education. This report uses data collected from two report forms each October: the Professional Assignment Information Form (PAIF) and the School Information Form (SIF).

For each public school in California, the PAIF contains individual-teacher-level and classroom-level data, which are later aggregated to the school level and weighted by full time equivalency. Variables at the individual level include gender, ethnicity, education level, experience, and types of certification held.[2] The PAIF also collects information on specific classes taught and student counts per section for each teacher. For the 1999–2000 school year, this dataset included over 330,000 individual observations and over 100 variables.

The school-level data in the SIF contain variables of two general types: (1) staff and student counts and (2) program types. The student counts are enumerated by grade and ethnicity. For simplification, we combine Filipinos, Asians, and Pacific Islanders into one category labeled "Asian." Program type variables identify information such as the extent of participation in the class size reduction program and the type of school

[1]Most of the information in this appendix is based on the discussion of the data sources in Appendix A of Betts, Rueben, and Danenberg (2000).

[2]The experience question asks for the number of years of service *including* the current school year. Given this definition, values of zero (which occur for a few observations) are recoded as missing.

(charter, alternative, etc.). The SIF contains over 500 variables for over 8,500 public schools in California. For 1997–1998 and 1999–2000, the school's community type (urban, suburban, or rural) was not included in the SIF. The data for this variable come from the 1996–1997 SIF and the Common Core of Data, a national dataset.

Free or Reduced-Price Lunch Enrollment

This school-level dataset contains counts and percentages of California children enrolled in free or reduced-price lunch programs and children in families receiving Aid to Families with Dependent Children (AFDC), which was replaced in 1996 by Temporary Assistance for Needy Families. According to the California Department of Education, schools report their meal program enrollment data annually, based on their October meal program enrollment files. AFDC data are collected each October through the cooperative efforts of the schools, districts, county offices of education, and county health and welfare offices. Both AFDC and subsidized-lunch data are collected on the California Department of Education Finance Division Form Number CFP-2 School Level AFDC Report.

The Language Census

The language census is a school-level summary that collects numerous types of data for elements in March of the current school year. This study uses the data on the number of LEP students—now known as EL students—and fluent English proficient students in California public schools by grade and by primary non-English language. From these counts, we calculate the percentages of Asian and Hispanic students who are ELs.

Standardized Test Scores

The file containing standardized test scores is maintained by the Standards, Curriculum, and Assessment Division of the California Department of Education. For the 1997–1998 and 1999–2000 school years, the file contains results from the *Stanford Achievement Test* Series, Ninth Edition, Form T (Stanford 9), administered by Harcourt, Brace &

Co. For each school, the test results are reported by grade in two ways for each subject: for all students tested in that grade, and for all LEP students tested in that grade. In 1999–2000, the file also contains results for non-LEP students. For 1997–1998, we calculate the test scores for the non-LEP students from the reported scores for the other two groups. Our analysis focuses on the mathematics and reading test score results for grades three and five, although the file contains results for other tests and for the 1998–1999 school year.

Appendix B

Additional Tables and Figures

This appendix contains additional information on student enrollment and teacher characteristics.

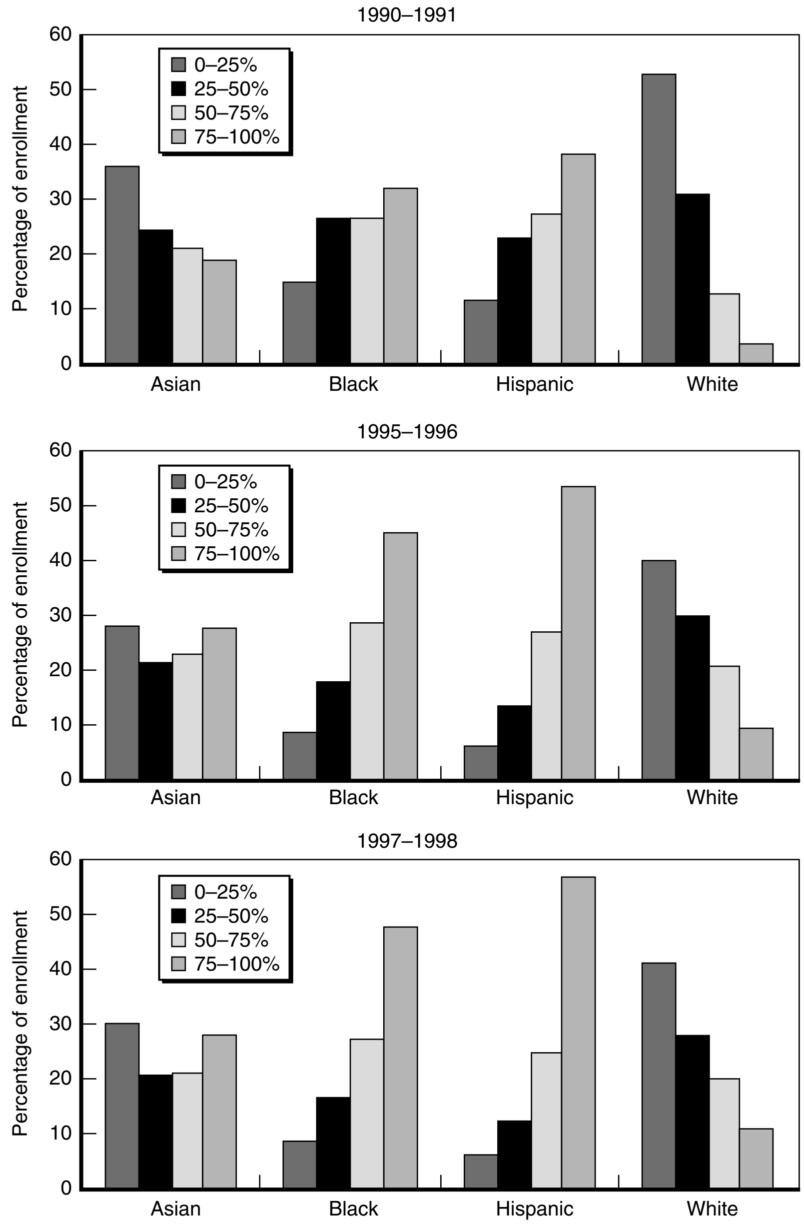


Figure B.1—Percentage Enrollment, by Percentage of Students Enrolled in Free or Reduced-Price Lunch Programs and Race/Ethnicity, 1990–1998

Table B.1

Average Class Size for Grades K–3, by Percentage of Students Enrolled in Free or Reduced-Price Lunch Programs and Race/Ethnicity, 1990–2000

% of Students in Lunch Program	1990–1991	1995–1996	1997–1998	1999–2000
Asian				
0–25	28.7	29.1	20.7	19.6
25–50	28.5	29.0	20.7	19.3
50–75	28.4	29.1	21.1	19.5
75–100	28.3	28.9	21.1	19.5
All	28.5	29.0	20.9	19.5
Black				
0–25	28.6	29.2	20.7	19.4
25–50	28.3	29.2	20.7	19.3
50–75	28.5	29.1	21.0	19.3
75–100	27.5	28.6	21.1	19.2
All	28.1	28.9	21.0	19.2
Hispanic				
0–25	28.5	29.0	20.7	19.3
25–50	28.5	29.3	20.9	19.3
50–75	28.5	29.3	21.2	19.5
75–100	27.7	29.1	21.5	19.4
All	28.2	29.2	21.3	19.4
White				
0–25	28.2	28.7	20.4	19.3
25–50	28.1	28.9	20.6	19.3
50–75	28.3	29.0	20.7	19.4
75–100	28.3	29.1	21.0	19.4
All	28.2	28.8	20.6	19.3
All Students				
	28.2	29.0	21.0	19.4

NOTE: The percentages are weighted by the number of students in each racial/ethnic category.

Table B.2

Average Class Size for Grades 4–5, by Percentage of Students Enrolled in Free or Reduced-Price Lunch Programs and Race/Ethnicity, 1990–2000

% of Students in Lunch Program	1990–1991	1995–1996	1997–1998	1999–2000
Asian				
0–25	29.7	30.3	30.1	29.9
25–50	29.7	30.2	30.1	29.9
50–75	29.6	30.0	29.5	29.6
75–100	29.1	29.5	29.6	29.5
All	29.6	30.0	29.8	29.7
Black				
0–25	29.5	30.3	30.0	29.9
25–50	29.3	30.2	29.7	29.8
50–75	29.2	29.8	29.2	29.5
75–100	28.1	29.6	29.4	29.2
All	28.9	29.8	29.5	29.5
Hispanic				
0–25	29.4	30.2	30.0	29.7
25–50	29.6	30.1	29.7	29.5
50–75	29.3	30.1	29.4	29.5
75–100	27.8	29.4	29.3	29.1
All	28.8	29.8	29.4	29.3
White				
0–25	29.3	29.9	29.6	29.4
25–50	29.3	29.8	29.5	29.5
50–75	29.1	29.8	29.3	29.4
75–100	28.9	29.7	29.2	29.2
All	29.3	29.9	29.5	29.4
All Students				
	29.1	29.8	29.5	29.4

NOTE: The percentages are weighted by the number of students in each racial/ethnic category.

Table B.3

Number of Elementary Schools, by Percentage of Students Enrolled in Free or Reduced-Price Lunch Programs, 1990–2000

% of Students in Lunch Program	1990–1991	1995–1996	1997–1998	1999–2000
All Schools				
0–25	1,546	1,088	1,099	1,192
25–50	1,209	1,015	964	974
50–75	854	1,096	1,068	1,074
75–100	661	1,236	1,402	1,420
All	4,270	4,435	4,533	4,660
Urban Schools				
0–25	263	174	230	218
25–50	401	217	227	255
50–75	415	374	363	367
75–100	462	770	834	814
All	1,541	1,535	1,654	1,654

Table B.4

Percentage of Students in Each Lunch Category, by Community Type and Race/Ethnicity, 1990–2000

	Asian				Black			
% of Students in Lunch Program	1990–1991	1995–1996	1997–1998	1999–2000	1990–1991	1995–1996	1997–1998	1999–2000
Urban								
0–25	18.3	13.7	17.8	19.6	7.5	4.0	4.6	4.5
25–50	23.4	14.7	16.0	20.6	22.0	10.2	11.0	12.4
50–75	28.3	24.6	20.5	21.7	26.3	26.4	24.2	25.9
75–100	29.8	47.1	45.6	38.2	44.0	59.4	60.2	57.2
Suburban								
0–25	54.9	41.1	43.3	45.9	28.3	16.1	16.3	19.4
25–50	24.0	27.1	24.5	24.1	33.0	29.4	25.4	25.1
50–75	13.2	21.5	20.9	19.3	26.1	31.5	31.4	29.3
75–100	7.6	10.3	11.2	10.8	12.4	22.9	26.9	26.3
Rural								
0–25	19.5	9.9	7.4	10.4	18.1	8.8	5.6	7.3
25–50	35.7	18.9	19.1	18.3	37.5	23.5	24.5	20.1
50–75	25.2	25.3	24.2	28.6	29.9	36.3	34.6	40.0
75–100	19.5	45.9	49.3	42.8	14.5	31.4	35.3	32.6

	Hispanic				White			
Urban								
0–25	4.9	2.3	2.7	2.6	31.6	28.2	32.1	35.0
25–50	14.6	5.8	6.2	7.6	35.5	24.5	23.7	23.3
50–75	24.8	18.8	18.1	17.3	23.3	26.3	24.0	22.1
75–100	55.6	73.0	72.9	72.5	9.3	20.9	20.1	19.6
Suburban								
0–25	23.5	12.6	12.1	13.2	66.4	50.5	52.7	56.6
25–50	33.7	22.6	19.7	19.6	24.8	28.4	26.1	24.0
50–75	27.9	34.9	32.9	31.2	7.3	16.4	15.1	13.7
75–100	14.8	29.9	35.2	36.0	1.3	4.7	6.1	5.7
Rural								
0–25	19.5	9.9	7.4	10.4	18.1	8.8	5.6	7.3
25–50	35.7	18.9	19.1	18.3	37.5	23.5	24.5	20.1
50–75	25.2	25.3	24.2	28.6	29.9	36.3	34.6	40.0
75–100	19.5	45.9	49.3	42.8	14.5	31.4	35.3	32.6

Table B.5

Teacher Experience, by Student Race/Ethnicity, Percentage of Students Enrolled in Free or Reduced-Price Lunch Programs, and Community Type, 1990–2000

	% with 0 Years Experience				% with 1 Year Experience			
% of Students in Lunch Program	1990–1991	1995–1996	1997–1998	1999–2000	1990–1991	1995–1996	1997–1998	1999–2000
Urban: Asian								
0–25	3.7	4.6	11.6	6.2	4.1	3.4	8.8	5.9
25–50	4.3	4.6	8.8	6.8	4.9	2.8	9.5	8.2
50–75	6.8	5.4	11.5	8.3	5.8	4.8	10.1	10.9
75–100	6.2	6.4	13.0	9.0	7.0	6.6	10.3	11.3
All	5.5	5.7	11.8	7.9	5.6	5.2	9.9	9.5
Urban: Black								
0–25	4.0	3.6	8.9	6.0	5.8	3.2	8.5	6.0
25–50	5.1	5.4	9.2	6.9	4.8	3.5	8.6	8.4
50–75	7.9	6.1	11.6	7.8	5.9	5.5	9.5	11.0
75–100	8.0	9.3	15.4	11.5	6.7	7.7	11.0	13.2
All	7.0	7.8	13.5	9.7	6.0	6.5	10.2	11.7
Urban: Hispanic								
0–25	5.9	4.9	10.7	6.9	5.4	3.8	8.2	6.2
25–50	5.2	6.0	10.8	6.8	4.9	4.4	8.6	7.2
50–75	7.1	7.1	13.2	7.9	6.7	5.1	8.9	10.3
75–100	6.9	8.0	15.1	10.4	6.9	6.6	9.5	12.3
All	6.6	7.6	14.4	9.6	6.5	6.1	9.3	11.4
Urban: White								
0–25	4.0	5.4	10.6	7.4	4.5	4.3	8.6	6.9
25–50	5.0	5.4	10.3	6.7	4.9	3.9	8.4	7.5
50–75	7.1	7.4	12.5	6.7	6.4	3.9	8.9	9.0
75–100	7.7	8.3	12.6	8.6	6.6	6.0	9.5	10.0
All	5.4	6.5	11.4	7.3	5.3	4.5	8.8	8.1
Urban: All Students								
	6.2	7.2	13.3	9.0	6.0	5.7	9.4	10.6
Suburban: Asian								
0–25	5.0	4.6	9.3	6.8	5.8	4.6	8.7	7.0
25–50	6.3	5.4	11.4	7.6	6.9	5.3	9.3	7.9
50–75	8.1	4.9	13.5	8.7	6.6	5.7	9.5	9.2
75–100	8.7	7.0	16.1	12.6	8.4	6.9	10.0	12.2
All	6.0	5.1	11.5	8.0	6.3	5.3	9.2	8.2

Table B.5 (continued)

% of Students in Lunch Program	% with 0 Years Experience 1990–1991	1995–1996	1997–1998	1999–2000	% with 1 Year Experience 1990–1991	1995–1996	1997–1998	1999–2000
Suburban: Black								
0–25	5.6	4.5	9.0	6.5	6.7	4.4	8.4	6.9
25–50	7.3	5.1	9.9	7.0	7.1	5.3	9.0	6.8
50–75	11.6	6.0	12.5	8.0	8.3	5.9	9.7	8.3
75–100	12.7	8.0	17.6	14.4	7.7	9.1	11.0	13.9
All	8.6	5.9	12.6	9.2	7.4	6.2	9.7	9.1
Suburban: Hispanic								
0–25	5.8	4.6	8.7	6.9	6.6	4.2	9.2	7.3
25–50	7.2	5.6	10.3	7.1	7.3	5.3	9.8	7.4
50–75	9.4	6.5	13.4	8.9	7.5	6.0	10.7	9.2
75–100	11.3	7.4	15.2	11.4	7.5	7.5	10.8	10.8
All	8.1	6.4	12.9	9.2	7.2	6.1	10.4	9.2
Suburban: White								
0–25	5.1	4.1	8.4	6.0	6.1	4.6	8.8	6.8
25–50	6.6	4.9	9.3	6.1	7.0	4.8	9.2	6.5
50–75	7.3	5.6	13.1	7.5	7.5	5.5	9.6	8.0
75–100	9.0	6.5	13.3	9.5	7.0	6.6	10.7	9.7
All	5.7	4.7	9.7	6.4	6.4	4.9	9.2	7.1
Suburban: All Students								
	6.6	5.4	11.2	7.8	6.7	5.4	9.6	8.1
Rural: Asian								
0–25	5.4	4.0	6.0	4.9	6.9	3.9	8.0	5.9
25–50	7.4	3.7	7.5	4.8	6.7	2.7	8.3	5.3
50–75	6.3	3.4	6.8	6.2	6.5	4.7	9.2	6.2
75–100	9.6	5.3	10.0	5.7	8.3	4.8	12.3	5.9
All	7.1	4.4	8.4	5.6	7.0	4.3	10.5	5.9
Rural: Black								
0–25	5.3	3.0	6.2	6.1	7.8	3.4	7.1	6.4
25–50	10.2	5.8	10.0	6.3	9.5	4.2	11.0	6.9
50–75	8.0	3.9	10.1	7.4	9.1	5.3	9.0	7.7
75–100	7.3	6.0	12.8	7.1	6.4	6.2	13.1	7.9
All	8.2	4.9	10.8	7.0	8.6	5.2	10.8	7.5
Rural: Hispanic								
0–25	5.9	4.4	6.7	3.8	8.1	3.5	7.7	6.0
25–50	6.7	4.5	8.3	5.1	7.2	3.3	8.8	5.9
50–75	7.9	4.0	9.0	6.6	8.6	5.2	9.9	7.7
75–100	8.3	5.7	13.5	8.6	7.2	6.1	12.5	8.9
All	7.5	4.9	11.2	7.3	7.8	5.2	11.0	8.0

Table B.5 (continued)

% of Students in Lunch Program	% with 0 Years Experience 1990–1991	1995–1996	1997–1998	1999–2000	% with 1 Year Experience 1990–1991	1995–1996	1997–1998	1999–2000
Rural: White								
0–25	4.5	4.0	5.5	2.8	6.5	3.5	7.9	4.8
25–50	6.7	3.8	6.7	4.1	7.0	2.9	8.3	4.9
50–75	7.4	4.0	7.8	5.6	7.7	4.4	9.3	6.1
75–100	9.7	4.8	9.7	5.8	6.5	5.2	11.3	6.0
All	6.1	4.0	7.2	4.5	6.9	3.7	8.9	5.4
Rural: All Students								
	6.7	4.4	9.0	5.8	7.2	4.3	9.9	6.6

NOTE: The percentages are weighted by the number of students in each racial/ethnic category.

Table B.6

Teacher Credentials and Education, by Student Race/Ethnicity, Percentage of Students Enrolled in Free or Reduced-Price Lunch Programs, and Community Type, 1990–2000

	% Not Fully Certified				% with Bachelor's Degree Only			
% of Students in Lunch Program	1990–1991	1995–1996	1997–1998	1999–2000	1990–1991	1995–1996	1997–1998	1999–2000
Urban: Asian								
0–25	0.1	0.3	6.1	9.8	10.7	13.4	19.8	18.7
25–50	0.1	0.6	7.4	10.5	14.1	11.9	17.7	15.5
50–75	0.2	1.4	12.4	17.6	20.3	16.5	24.8	23.3
75–100	0.2	2.5	15.9	19.8	25.6	20.8	27.0	32.4
All	0.2	1.7	12.1	15.5	18.7	17.4	23.8	24.3
Urban: Black								
0–25	0.5	0.5	6.9	12.9	11.2	11.4	17.7	18.4
25–50	0.7	1.0	10.1	11.2	15.6	15.2	21.8	19.9
50–75	0.6	2.7	14.6	20.4	20.0	16.0	25.3	25.4
75–100	0.2	5.3	26.8	32.0	32.2	26.7	34.7	38.9
All	0.4	4.0	21.1	25.5	23.7	22.1	30.3	32.1
Urban: Hispanic								
0–25	0.2	0.4	6.5	8.5	12.8	15.5	20.9	19.9
25–50	0.4	1.0	10.0	10.3	16.3	16.0	22.3	18.0
50–75	0.8	2.6	16.1	17.9	23.0	17.8	29.2	27.3
75–100	0.2	3.9	25.5	30.2	34.6	29.8	37.8	40.9
All	0.4	3.4	22.3	26.0	28.0	26.4	34.8	36.2
Urban: White								
0–25	0.1	0.4	5.5	7.2	11.0	18.0	17.3	17.1
25–50	0.2	0.8	8.4	9.3	16.2	14.9	18.4	18.2
50–75	0.4	1.4	11.5	13.4	22.1	15.6	22.1	22.6
75–100	0.2	2.0	15.5	17.7	27.2	21.0	27.9	28.7
All	0.2	1.1	9.7	11.1	16.9	17.2	20.9	20.8
Urban: All Students								
	0.3	2.8	18.4	21.9	23.3	22.8	30.1	31.4
Suburban: Asian								
0–25	0.1	0.4	5.3	5.3	13.4	7.9	12.4	12.0
25–50	0.4	0.5	6.1	7.9	14.9	13.5	14.8	16.6
50–75	0.3	0.7	8.1	8.6	18.7	14.1	16.4	19.3
75–100	0.1	2.1	11.9	16.1	26.7	24.4	22.1	27.5
All	0.2	0.6	6.9	7.7	15.4	12.4	15.0	16.2

Table B.6 (continued)

% of Students in Lunch Program	% Not Fully Certified 1990–1991	1995–1996	1997–1998	1999–2000	% with Bachelor's Degree Only 1990–1991	1995–1996	1997–1998	1999–2000
Suburban: Black								
0–25	0.2	0.5	5.5	5.6	13.3	11.0	14.6	12.3
25–50	0.7	0.6	6.0	8.3	16.2	13.9	16.4	15.0
50–75	2.4	1.1	8.8	9.0	26.1	14.6	18.3	18.9
75–100	1.7	2.7	12.0	16.4	32.8	31.5	26.2	35.1
All	1.1	1.2	8.5	10.1	20.1	17.7	19.3	20.9
Suburban: Hispanic								
0–25	0.2	0.4	5.6	5.9	13.3	8.1	12.1	12.2
25–50	0.8	0.8	7.6	8.5	15.2	11.6	14.5	15.0
50–75	1.2	1.5	12.3	12.4	18.6	12.2	17.8	18.7
75–100	0.9	3	15.2	18.4	25.2	19.3	22.4	26.7
All	0.8	1.6	11.6	13	17.2	13.7	18.1	20.0
Suburban: White								
0–25	0.1	0.3	4.5	4.0	11.9	8.1	11.6	11.5
25–50	0.3	0.5	6.0	6.1	14.6	11.4	12.6	13.6
50–75	0.2	0.8	8.1	8.7	17.6	12.4	16.3	17.5
75–100	0.5	1.5	10.4	11.7	20.9	17.6	19.0	21.0
All	0.1	0.5	5.8	5.6	13.1	10.2	13.0	13.3
Suburban: All Students								
	0.4	0.9	8.1	8.9	14.9	12.1	15.5	16.7
Rural: Asian								
0–25	0.2	0.2	3.1	5.3	11.8	6.8	8.8	9.9
25–50	0.5	0.6	5.6	5.8	15.4	9.5	13.3	12.4
50–75	0.3	0.3	5.5	6.4	12.6	8.2	12.1	12.5
75–100	0.1	0.4	7.9	10.2	14.9	9.6	17.5	14.2
All	0.3	0.4	6.5	7.8	13.9	8.9	14.8	12.9
Rural: Black								
0–25	0.2	0.1	3.3	6.8	14.1	8.0	8.9	9.5
25–50	1.3	1.4	8.9	9.6	24.6	14.0	19.1	18.1
50–75	0.7	0.7	8.5	10.2	18.5	11.7	18.3	20.1
75–100	0.1	1.6	12.3	14.9	14.0	10.1	19.3	16.7
All	0.7	1.1	9.6	11.4	19.3	11.4	18.3	17.8
Rural: Hispanic								
0–25	0.3	0.2	3.9	4.5	14.3	7.7	10.4	9.6
25–50	0.8	1	6.9	6.8	15.1	9.4	13.9	12.1
50–75	0.6	0.8	8	10.1	14.9	8.0	13.3	15.0
75–100	0.6	1.6	13.3	17.2	14.3	11.1	20.5	20.3
All	0.6	1.1	10.5	13.1	14.7	9.7	17.1	17.2

Table B.6 (continued)

% of Students in Lunch Program	% Not Fully Certified 1990–1991	1995–1996	1997–1998	1999–2000	% with Bachelor's Degree Only 1990–1991	1995–1996	1997–1998	1999–2000
Rural: White								
0–25	0.2	0.1	2.9	3.0	11.1	7.2	8.8	8.5
25–50	0.7	0.5	4.2	4.6	14.8	8.2	11.9	10.6
50–75	0.5	0.6	6.1	7.0	14.6	9.0	13.1	14.2
75–100	0.3	0.9	7.0	9.1	11.1	8.6	14.8	15.1
All	0.4	0.5	4.9	5.5	13.3	8.3	12.1	11.8
Rural: All Students								
	0.5	0.8	7.4	9.1	13.9	8.9	14.4	14.4

NOTE: The percentages are weighted by the number of students in each racial/ethnic category.

Appendix C

Empirical Model for Student Achievement

This appendix describes the empirical model used in the analysis of student achievement in Chapter 4. It also includes the tables with the results of these regressions.

Our analysis focuses on CSR in third grade only. Implementation in first and second grades was nearly universal, and California does not test students in kindergarten. Furthermore, individual-level data from year to year are not available, so it is not possible to follow students from kindergarten to later years. We use data from 1997–1998 and 1999–2000 to estimate the following model:

$$A_{sc} = X_{sc}\alpha_x + CS_{sc}\alpha_{cs} + TC_{sc}\alpha_{TC} + I_{imp} * d_c + d_c + \theta_x + \omega_{sc} + \varepsilon_{sc} \quad (1)$$

Equation (1) describes third-grade student achievement (A_{sc}) for school s and cohort c. The vector TC includes the percentage of teachers in their first year, the percentage in their second year, the percentage not fully certified, and the percentage with no graduate education. The vector X includes percentage black, percentage Hispanic, percentage Asian, percentage enrolled in subsidized lunch programs, and percentage EL.[1] A dummy variable (d_c) equal to 1 for the 1999–2000 cohort captures any upward drift in the test score as well as the effects of other state policies and systematic changes over time. Because the schools that implemented CSR in 1998–1999 or 1999–2000 are likely different from other schools, the equation contains a dummy variable for these schools (I_{imp}). This variable is interacted with d_c so that it controls for any

[1]Because of data constraints, we use school-level averages for percentage enrolled in subsidized lunch programs and percentage EL. All other variables are available at the grade level.

systematic differences in the rate of test score growth over time between early and late adopters. Unobserved school quality is decomposed into a permanent (θ) and a time-varying component (ω), the latter of which primarily reflects changes over time in teacher quality that are not captured by the included characteristics.

Ordinary least squares estimation of Eq. (1) that does not control for unobserved differences among schools is probably going to generate biased estimates, because there are many differences among schools not captured by the limited set of included variables. Therefore, we include a fixed effect for each school that captures all systematic differences among schools that are stable over time. The estimates are identified by within-school changes in class size, teacher characteristics, and achievement.

It is important to note that the school fixed effects do not control for other factors that may change over time at schools and that may be correlated with class size. One such factor is teacher quality. If teacher quality tends to decline as class size is reduced because schools hire lower-quality teachers than those already teaching (and the decline is not completely captured by the included teacher variables), the estimated class size effect will combine such a change in teacher quality with the direct effect of smaller classes.

In most cases, this would present a serious problem, but we believe that the identification of any change in teacher quality that accompanied CSR is crucial to a comprehensive understanding of policy effectiveness. This is very difficult in the case of CSR, because many schools had fully implemented CSR before the first year of data were collected, and the remaining schools likely differ in important ways. Nevertheless, we attempt to learn something about the decline in instructional effectiveness by comparing class size effects for schools that adopted CSR during our sample to those for schools that adopted CSR earlier (or never).

Variation in class size for early adopters results entirely from small year-to-year changes in enrollment that lead class size to fluctuate at or below 20 students.[2] Once schools have implemented CSR, there is no

---

[2]Similarly, schools that did not adopt CSR as of 1999–2000 had variation in class size based on small year-to-year changes in enrollment.

longer a need to hire additional teachers, and these fluctuations are unlikely to be correlated with systematic changes in teacher quality. On the other hand, the pre-CSR/post-CSR difference in class size is driven predominantly by the hiring of additional teachers. Therefore, class-size estimates from these schools are likely to capture any decline in teacher quality that is linked to the magnitude of the reduction in class size needed to satisfy CSR. Note that the more class size had to be reduced the larger is the proportional increase in the teaching force.

This technique does not capture all of the changes in CSR that occurred throughout the state. First, teacher quality declines resulting from the hiring away of teachers by other districts, even after the school in question has reduced class size, are not captured. At the same time, there is concern that other factors (e.g., facility difficulties) may reduce the benefits of smaller classes for late adopters. The late-adopter fixed effect captures systematic differences in the change in test scores between 1997 and 1999, so the class-size estimates would capture only such differences that are correlated with the change in class size.

To separate short- from long-term effects, we actually estimate two equations, with and without teacher characteristics:

$$\begin{aligned} A_{sc} = {} & X_{sc}\alpha_x + CS_{sc} * I_{imp}\alpha_{CSimp} + CS_{sc} * I_{not}\alpha_{CSnot} \\ & + \theta_s + d_c + d_c * I_{imp} + \omega_{sc} + \varepsilon_{sc} \end{aligned} \tag{2}$$

$$\begin{aligned} A_{sc} = {} & X_{sc}\alpha_x + CS_{sc} * I_{imp}\alpha_{CSimp} + CS_{sc} * I_{not}\alpha_{CSnot} \\ & + TC_{sc} * I_{imp}\alpha_{TCimp} + TC_{sc} * I_{not}\alpha_{TCnot} \\ & + \theta_s + d_c + d_c * I_{imp} + \omega_{sc} + \varepsilon_{sc} \end{aligned} \tag{3}$$

I_{not} is a dummy variable for schools that did not implement CSR in 1998–1999 or 1999–2000. All other variables (including I_{imp}) are defined as in Eq. (1).

How should the results be interpreted? We predict that the need to hire large numbers of inexperienced teachers reduces the class-size effect for schools implementing CSR late ($I_{imp} = 1$), so that it is smaller than the effect for other schools. One possibility is that a smaller effect for late adopters reflects the temporary problems at the time of CSR implementation faced by many schools, including the addition of many new teachers. Under this interpretation, other aspects of schools including teacher quality will return to pre-CSR levels after a few years. If the inclusion of teacher experience, certification, and education eliminates much or all of the gap in class size effects between early and late adopters, it would suggest that the gap reflects a temporary cost.

An alternative possibility is that the difference between early and late adopters captures a permanent decline in teacher quality that moves affected schools to a lower long-run average. If the inclusion of teacher experience, certification, and education eliminates little or none of the gap, it would suggest that the gap reflects longer-term reductions in teacher quality that likely affected all relevant schools. Because the effect of CSR may differ by student demographics, we allow the effects of class size to vary along a number of dimensions.

All in all, this approach attempts to uncover a portion of the change in teacher quality associated with CSR implementation. The teacher characteristic tables in Chapter 3 provide clear evidence that CSR led many schools (particularly those with predominantly poor and minority student bodies) to hire not fully certified and inexperienced teachers. Unfortunately, the included teacher characteristics measure only some aspects of teacher quality, which leads us to an alternative approach based on differences in effect sizes by timing of CSR implementation. This approach is likely to provide additional information, although, as noted above, it does not capture any changes in teacher quality not related to the school's own reduction of class size.

Table C.1

Estimated Effects of Class Size and Teacher Characteristics on the Percentage of Students Exceeding the National Median Test Scores

	Mathematics		Reading	
	No	Yes	No	Yes
Class size	–0.40	–0.41	–0.26	–0.26
	(4.85)	(5.01)	(3.85)	(3.99)
Teacher characteristics				
Percentage with 0 years experience		–0.027		–0.030
		(2.43)		(3.24)
Percentage with 1 year experience		–0.0062		–0.0031
		(0.56)		(0.34)
Percentage not fully certified		–0.018		–0.011
		(1.38)		(1.07)
Percentage with bachelor's degree only		–0.0034		0.0094
		(0.33)		(1.14)
No. of observations	7,626	7,626	7,612	7,612

NOTES: Absolute value of t-statistics are in parentheses. The dependent variable is the school percentage of non-EL students who exceed the 50th percentile on a nationally determined basis. Other regressors include percentage black, percentage Hispanic, percentage Asian, percentage of students enrolled in free or reduced-price lunch programs, percentage EL, a dummy variable for the year 1999, and the 1999 dummy variable interacted with a dummy variable for schools adopting CSR in 1998 or 1999. All regressions are weighted by the number of test-takers. Column headings labeled yes or no indicate whether teacher characteristics are included.

Table C.2

Estimated Effects of Class Size and Teacher Characteristics on Third-Grade Test Scores: Robustness Analysis for Credential and Experience

	Mathematics				Reading			
Class size	–0.38	–0.41	–0.41	–0.41	–0.16	–0.27	–0.26	–0.26
	(7.50)	(5.02)	(4.99)	(5.00)	(3.92)	(4.04)	(3.97)	(3.97)
Teacher characteristics								
Percentage with 0 years experience			–0.030	–0.027			–0.032	–0.029
			(2.72)	(2.41)			(3.58)	(3.21)
Percentage with 1 year experience			–0.008	–0.006			–0.005	–0.003
			(0.70)	(0.55)			(0.51)	(0.33)
Average experience		0.067				0.101		
		(1.54)				(2.88)		
Percentage not fully certified	–0.029	–0.025			–0.022	–0.017		
	(2.27)	(1.96)			(2.15)	(1.61)		
Percentage intern				–0.029				–0.028
				(1.24)				(1.48)
Percentage with emergency certification or waiver			–0.011	–0.016			–0.003	–0.007
			(0.82)	(1.10)			(0.28)	(0.63)
Percentage with bachelor's degree only	–0.008	–0.005	–0.005	–0.003	0.005	0.009	0.007	0.009
	(0.82)	(0.49)	(0.54)	(0.33)	(0.60)	(1.15)	(0.91)	(1.14)
No. of observations	7,626	7,626	7,626	7,626	7,612	7,612	7,612	7,612

NOTES: Absolute value of t-statistics are in parentheses. The dependent variable is the school percentage of non-EL students who exceed the 50th percentile on a nationally determined basis. Each column includes only a subset of the listed teacher characteristics, identified by the presence of estimated coefficients and t-statistics. For example, the first column contains results from a model that includes only two teacher characteristics: the percentage of teachers not fully certified and the percentage with a bachelor's degree. Other regressors include percentage black, percentage Hispanic, percentage Asian, percentage of students enrolled in free or reduced-price lunch programs, percentage EL, a dummy variable for the year 1999, and the 1999 dummy variable interacted with a dummy variable for schools adopting CSR in 1998 or 1999.

Table C.3

Estimated Effects of Class Size on Third-Grade Test Scores, by Community Type

	Urban		Suburban		Rural		Los Angeles Unified		Five Next Largest Districts Combined	
	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Mathematics										
Class size	–0.473	–0.480	–0.374	–0.394	–0.248	–0.230	–0.183	–0.172	–1.083	–1.054
	(3.52)	(3.57)	(3.18)	(3.34)	(1.07)	(0.99)	(0.62)	(0.58)	(3.40)	(3.26)
No. of observations	2,966	2,966	3,494	3,494	1,166	1,166	816	816	644	644
Reading										
Class size	–0.380	–0.385	–0.214	–0.222	–0.095	–0.070	–0.176	–0.149	–0.882	–0.837
	(3.41)	(3.46)	(2.32)	(2.42)	(0.49)	(0.36)	(0.69)	(0.58)	(3.30)	(3.09)
No. of observations	2,972	2,972	3,484	3,484	1,156	1,156	816	816	646	646

NOTES: Absolute value of t-statistics are in parentheses. The dependent variable is the school percentage of non-EL students who exceed the 50th percentile on a nationally determined basis. Other regressors include percentage black, percentage Hispanic, percentage Asian, percentage of students enrolled in free or reduced-price lunch programs, percentage EL, a dummy variable for the year 1999, and the 1999 dummy variable interacted with a dummy variable for schools adopting CSR in 1998 or 1999. The models with teacher characteristics include additional controls for experience, certification, and education. Column headings labeled yes or no indicate whether teacher characteristics are included.

Table C.4

Estimated Effects of Class Size, by Student Demographic Composition and Community Type

	All Schools		Urban		Suburban		Rural		L. A. Unified		5 Next Largest Districts Combined	
	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Mathematics												
Class size	–0.398	–0.411	–0.473	–0.480	–0.374	–0.394	–0.248	–0.230	–0.183	–0.172	–1.083	–1.054
	(4.85)	(5.01)	(3.52)	(3.57)	(3.18)	(3.34)	(1.07)	(0.99)	(0.62)	(0.58)	(3.40)	(3.26)
Class size	–0.175	–0.207	–0.246	–0.280	–0.029	–0.089	–0.476	–0.423	–0.921	–1.261	–0.081	–0.061
	(1.71)	(2.02)	(1.32)	(1.49)	(0.20)	(0.63)	(1.51)	(1.31)	(1.57)	(2.15)	(0.14)	(0.10)
Class size interaction	–0.005	–0.004	–0.004	–0.004	–0.008	–0.007	0.004	0.003	0.010	0.013	–0.014	–0.012
with. % free lunch	(3.66)	(3.34)	(1.74)	(1.62)	(4.50)	(3.91)	(1.06)	(0.75)	(1.46)	(1.95)	(2.10)	(1.74)
Class size	–0.462	–0.468	–0.585	–0.573	–0.378	–0.396	–0.338	–0.371	–0.537	–0.504	–0.944	–0.856
	(5.36)	(5.43)	(4.10)	(4.01)	(3.03)	(3.18)	(1.37)	(1.48)	(1.73)	(1.60)	(2.27)	(2.01)
Class size interaction	0.007	0.006	0.009	0.008	0.000	0.001	0.017	0.025	0.020	0.020	–0.005	–0.008
with. % black	(2.41)	(2.25)	(2.34)	(1.99)	(0.09)	(0.22)	(1.01)	(1.42)	(3.43)	(3.11)	(0.52)	(0.72)
No. of observations	7,626	7,626	2,966	2,966	3,494	3,494	1,166	1,166	816	816	644	644
Reading												
Class size	–0.255	–0.264	–0.380	–0.385	–0.214	–0.222	–0.095	–0.070	–0.176	–0.149	–0.882	–0.837
	(3.85)	(3.99)	(3.41)	(3.46)	(2.32)	(2.42)	(0.49)	(0.36)	(0.69)	(0.58)	(3.30)	(3.09)
Class size	–0.102	–0.127	–0.234	–0.268	0.024	0.006	–0.148	–0.117	–0.875	–1.010	–0.661	–0.377
	(1.24)	(1.53)	(1.51)	(1.72)	(0.22)	(0.05)	(0.55)	(0.43)	(1.73)	(1.97)	(1.37)	(0.72)
Class size interaction	–0.003	–0.003	–0.002	–0.002	–0.005	–0.005	0.001	0.000	0.009	0.011	–0.003	–0.004
with. % free lunch	(3.11)	(2.88)	(1.37)	(1.23)	(3.95)	(3.73)	(0.28)	(0.02)	(1.60)	(1.83)	(0.55)	(0.71)
Class size	–0.242	–0.245	–0.373	–0.365	–0.197	–0.202	–0.059	–0.044	–0.305	–0.208	–0.532	–0.386
	(3.48)	(3.53)	(3.15)	(3.08)	(2.02)	(2.08)	(0.29)	(0.21)	(1.13)	(0.76)	(1.53)	(1.08)
Class size interaction	–0.001	–0.002	–0.001	–0.002	–0.002	–0.002	–0.007	–0.004	0.007	0.004	–0.013	–0.015
with. % black	(0.62)	(0.86)	(0.18)	(0.64)	(0.53)	(0.49)	(0.48)	(0.27)	(1.42)	(0.77)	(1.57)	(1.76)
No. of observations	7,612	7,612	2,972	2,972	3,484	3,484	1,156	1,156	816	816	646	646

NOTES: Absolute value of t-statistics are in parentheses. The dependent variable is the percentage of third-grade students who exceed the 50th percentile on a nationally determined basis. All regressions include school fixed effects, use data from 1997 and 1999, and are weighted by the number of test-takers. Other regressors include percentage black, percentage Hispanic, percentage Asian, percentage of students enrolled in free or reduced-price lunch programs, percentage EL, a dummy variable for the year 1999, and the 1999 dummy variable interacted with a dummy variable for schools adopting CSR in 1998 or 1999. The models with teacher characteristics include additional controls for experience, certification, and education. The column headings labeled yes or no indicate whether teacher characteristics are included.

Table C.5

Estimated Effects of Class Size and Teacher Characteristics on Third-Grade Test Scores Combining All Student Demographic Interactions, by Teacher Characteristics

	Mathematics		Reading	
	No	Yes	No	Yes
Class size	–0.212	–0.244	–0.104	–0.127
	(2.07)	(2.37)	(1.25)	(1.52)
Class size interaction with % free lunch	–0.006	–0.005	–0.003	–0.003
	(4.47)	(4.02)	(3.06)	(2.71)
Class size interaction with % black	0.010	0.009	0.001	0.000
	(3.52)	(3.12)	(0.22)	(0.17)
No. of observations	7,626	7,626	7,612	7,612

NOTES: Absolute value of t-statistics are in parentheses. The dependent variable is the school percentage of non-EL students who exceed the 50th percentile on a nationally determined basis. Other regressors include percentage black, percentage Hispanic, percentage Asian, percentage of students enrolled in free or reduced-price lunch programs, percentage EL, a dummy variable for the year 1999, and the 1999 dummy variable interacted with a dummy variable for schools adopting CSR in 1998 or 1999. The models with teacher characteristics include additional controls for experience, certification, and education. Column headings labeled yes or no indicate whether teacher characteristics are included.

Table C.6

Variable Means and Standard Deviations for Third Grade, by Timing of CSR: Change in Variable from 1997 to 1999 Below Overall Statistic

	Did Not Change CSR		Changed CSR	
	Mean	S. D.	Mean	S. D.
Math test score	52.5	23.5	46.8	22.3
Reading test score	47.0	22.6	40.9	20.6
Class size	20.5	3.7	23.4	5.0
	–0.2	2.2	–8.7	3.2
Teacher characteristics				
Percentage with 0	9.1	15.7	10.1	17.0
years experience	–6.9	20.8	–6.8	22.8
Percentage with 1	8.1	14.2	9.2	15.0
year experience	–2.6	19.7	1.3	20.9
Percentage not	11.5	19.0	15.5	21.3
fully certified	–1.6	19.2	0.1	21.0
Percentage with bachelor's	21.2	26.1	24.2	26.5
degree only	–1.7	21.9	–0.5	25.9
No. of observations	4,748		2,864	

NOTE: All statistics are weighted by the number of students in the grade.

Table C.7

Estimated Effects of Class Size on Third-Grade Test Scores, by Timing of CSR, Teacher Characteristics, and Student Demographic Composition: Separate Regressions for Each Column and Panel

	Mathematics		Reading	
	No	Yes	No	Yes
Class size	–0.366	–0.377	–0.315	–0.321
Did not change CSR	(2.85)	(2.94)	(3.04)	(3.10)
Class size	–0.421	–0.433	–0.214	–0.224
Changed CSR	(3.95)	(4.06)	(2.49)	(2.60)
Class size	–0.002	–0.059	–0.021	–0.065
Did not change CSR	(0.01)	(0.32)	(0.14)	(0.43)
Class size	–0.219	–0.209	–0.083	–0.097
Changed CSR	(1.81)	(1.71)	(0.85)	(0.98)
Class size interaction with % free lunch	–0.007	–0.006	–0.005	–0.005
Did not change CSR	(2.76)	(2.32)	(2.76)	(2.34)
Class size interaction with % free lunch	–0.004	–0.005	–0.003	–0.003
Changed CSR	(3.50)	(3.67)	(2.84)	(2.76)
Class size	–0.379	–0.394	–0.245	–0.252
Did not change CSR	(2.77)	(2.88)	(2.22)	(2.28)
Class size	–0.476	–0.458	–0.194	–0.196
Changed CSR	(4.34)	(4.17)	(2.19)	(2.20)
Class size interaction with % black	0.001	0.002	–0.008	–0.008
Did not change CSR	(0.12)	(0.33)	(1.86)	(1.82)
Class size interaction with % black	0.007	0.005	–0.001	–0.002
Changed CSR	(2.42)	(1.61)	(0.58)	(0.86)
No. of observations	7,626	7,626	7,612	7,612

NOTES: Absolute value of t-statistics are in parentheses. The dependent variable is the percentage of third-grade students who exceed the 50th percentile on a nationally determined basis. All regressions include school fixed effects, use data from 1997 and 1999, and are weighted by the number of test-takers. Other regressors include percentage black, percentage Hispanic, percentage Asian, percentage of students enrolled in free or reduced-price lunch programs, percentage EL, and a dummy variable for the year 1999. The models with teacher characteristics include additional controls for experience, certification, and education. Column headings labeled yes or no indicate whether teacher characteristics are included.

Table C.8

Estimated Effects of Class Size on Third-Grade Test Scores Combining All Interactions, by Timing of CSR, Teacher Characteristics, and Student Demographic Composition

	Mathematics		Reading	
	No	Yes	No	Yes
Class size Did not change CSR	0.018 (0.10)	–0.038 (0.20)	0.019 (0.13)	–0.036 (0.24)
Class size Changed CSR	–0.251 (2.06)	–0.227 (1.84)	–0.076 (0.77)	–0.089 (0.89)
Class size interaction with % free lunch Did not change CSR	–0.008 (3.23)	–0.007 (2.89)	–0.005 (2.65)	–0.004 (2.13)
Class size interaction with % free lunch Changed CSR	–0.006 (4.35)	–0.006 (4.25)	–0.003 (2.85)	–0.003 (2.65)
Class size interaction with % black Did not change CSR	0.004 (0.80)	0.005 (0.96)	–0.006 (1.32)	–0.006 (1.35)
Class size interaction with % black Changed CSR	0.011 (3.58)	0.008 (2.75)	0.001 (0.30)	0.0001 (0.05)
No. of observations	7,626	7,626	7,612	7,612

NOTES: Absolute value of t-statistics are in parentheses. The dependent variable is the percentage of third-grade students who exceed the 50th percentile on a nationally determined basis. All regressions include school fixed effects, use data from 1997 and 1999, and are weighted by the number of test-takers. Other regressors include percentage black, percentage Hispanic, percentage Asian, percentage of students enrolled in free or reduced-price lunch programs, percentage EL, a dummy variable for the year 1999, and the 1999 dummy variable interacted with a dummy variable for schools adopting CSR in 1998 or 1999. The models with teacher characteristics include additional controls for experience, certification, and education. Column headings labeled yes or no indicate whether teacher characteristics are included.

Table C.9

Estimated Effects of Third-Grade and Fifth-Grade Class Size on Fifth-Grade Test Scores

	Mathematics				Reading			
	No	Yes	No	Yes	No	Yes	No	Yes
Class size fifth grade	–0.244	–0.259	–0.263	–0.268	–0.071	–0.081	–0.079	–0.080
	(4.00)	(4.23)	(3.39)	(3.44)	(1.35)	(1.54)	(1.19)	(1.19)
Class size third grade	–0.0005	–0.0084	–0.0332	–0.0303	0.0297	0.0252	–0.0025	–0.0004
	(0.01)	(0.20)	(0.75)	(0.68)	(0.83)	(0.71)	(0.07)	(0.01)
Class size fifth grade interaction with % black			0.0018	0.0003			0.0007	–0.0009
			(0.38)	(0.07)			(0.18)	(0.21)
Class size third grade interaction with % black			0.0038	0.0029			0.0038	0.0033
			(2.10)	(1.55)			(2.44)	(2.08)
No. of observations	6,750	6,750	6,750	6,750	6,718	6,718	6,718	6,718

NOTES: Absolute value of t-statistics are in parentheses. The dependent variable is the percentage of fifth-grade students who exceed the 50th percentile on a nationally determined basis. All regressions include school fixed effects, use data from 1997 and 1999, and are weighted by the number of test-takers. Other regressors include percentage black, percentage Hispanic, percentage Asian, percentage of students enrolled in free or reduced-price lunch programs, percentage EL, and a dummy variable for the year 1999. Column headings labeled yes or no indicate whether teacher characteristics are included.

Table C.10

Variable Means and Standard Deviations for Fifth Grade, by Timing of CSR: Change in Variable from 1997 to 1999 Below Overall Statistic

	Did Not Change CSR		Changed CSR	
	Mean	S. D.	Mean	S. D.
Math test score	52.2	22.1	47.0	20.4
Reading test score	49.5	21.5	43.7	19.4
Class size	29.4	3.2	30.1	2.6
	0.0	3.7	–0.2	2.8
Teacher characteristics				
Percentage with 0	9.8	19.6	11.9	20.2
years experience	–1.0	26.5	–0.1	27.4
Percentage with 1	8.5	17.9	9.3	18.2
year experience	1.5	25.0	4.0	25.3
Percentage not	11.9	22.4	14.5	22.5
fully certified	3.9	23.0	6.3	25.2
Percentage with bachelor's degree only	20.5	28.6	23.6	28.5
	2.9	26.6	5.8	29.8
No. of observations	4,634		2,818	

NOTE: All statistics are weighted by number of students in the grade.

References

Angrist, Joshua D., and Victor Lavy, "Using Maimonides' Rule to Estimate the Effect of Class Size on Scholastic Achievement," *Quarterly Journal of Economics*, Vol. 114, No. 2, 1999, pp. 533–576.

Betts, Julian R., Kim S. Rueben, and Anne Danenberg, *Equal Resources, Equal Outcomes? The Distribution of School Resources and Student Achievement in California*, Public Policy Institute of California, San Francisco, California, 2000.

Bohrnstedt, George W., and Brian M. Stecher, eds., *Class Size Reduction in California: Early Evaluation Findings, 1996–1998*, American Institutes for Research, Palo Alto, California, 1999.

Bohrnstedt, George W., and Brian M. Stecher, eds., *Class Size Reduction in California: The 1998–1999 Evaluation Findings,* California Department of Education, Sacramento, California, 2000.

Bohrnstedt, George W., and Brian M. Stecher, eds., *Class Size Reduction in California: Findings from 1999–00 and 2000–01*, California Department of Education, Sacramento, California, 2002.

Ehrenberg, Ronald G., and Dominic Brewer, "Do School and Teacher Characteristics Matter? Evidence from High School and Beyond," *Economics of Education Review*, Vol. 13, No. 1, 1994, pp. 1–17.

Ehrenberg, Ronald G., Dominic Brewer, Adam Gamoran, and J. Douglas Willms, "Class Size and Student Achievement," *Psychological Science in the Public Interest*, Vol. 2, No. 1, 2001, pp. 1–30.

Ferguson, Ronald F., "Teachers' Perceptions and Expectations and the Black-White Test Score Gap," in Christopher Jencks and Meredith Phillips, eds., *The Black-White Test Score Gap*, The Brookings Institution Press, Washington, D.C., 1998, pp. 273–317.

Fetler, Mark, "High School Staff Characteristics and Mathematics Test Results," *Education Policy Analysis Archives*, Vol. 7, No. 9, 1999.

Goldhaber, Dan D., and Dominic J. Brewer, "Does Teacher Certification Matter? High School Teacher Certification Status and Student Achievement," *Educational Evaluation and Policy Analysis*, Vol. 22, No. 2, 2000, pp. 129–145.

Hanushek, Eric A., Steven G. Rivkin, and John F. Kain, "Do Higher Salaries Buy Better Teachers?" NBER Working Paper Number 7082, Cambridge, Massachusetts, 1999.

Hoxby, Caroline Minter, "The Effects of Class Size on Student Achievement: New Evidence from Population Variation," *Quarterly Journal of Economics*, Vol. 115, No. 4, 2000, pp. 1239–1286.

Krueger, Alan B., "Experimental Estimates of Education Production Functions," *Quarterly Journal of Economics*, Vol. 114, No. 2, 1999, pp. 497–532.

Krueger, Alan B., and Diane Whitmore, "The Effect of Attending a Small Class in the Early Grades on College-Test Taking and Middle School Test Results: Evidence from Project Star," *The Economic Journal*, Vol. 111, No. 2, 2001, pp. 1–28.

Lankford, Hamilton, Susanna Loeb, and James Wyckoff, "Teacher Preferences and the Plight of Urban Schools: A Descriptive Analysis of the New York State Teaching Force," Working Paper, SUNY, Albany, New York, 2001.

Lazear, Edward P., "Educational Production," *Quarterly Journal of Economics*, Vol. 116, No. 3, 2001, pp. 777–804.

Raymond, Margaret, Stephen Fletcher, and Javier Luque, "Teach for America: An Evaluation of Teacher Differences and Student Outcomes in Houston, Texas," Working Paper, Stanford University, Stanford, California, 2001.

Reichardt, Robert E., *The Cost of Class Size Reduction: Advice for Policymakers*, Ph.D. Dissertation, The RAND Graduate School, Santa Monica, California, 2000.

Rivkin, Steven G., Eric A. Hanushek, and John F. Kain, "Teachers, Schools, and Academic Achievement," Working Paper, University of Texas, Dallas, Texas, 2000.

Ross, Randy, "How Class-Size Reduction Harms Kids in Poor Neighborhoods," *Education Week*, May 26, 1999.

Sanders, William L., and Sandra P. Horn, "The Tennessee Value-Added Assessment System (TVAAS): Mixed-Model Methodology in Educational Assessment," *Journal of Personnel Evaluation in Education*, Vol. 8, No. 2, 1994, pp. 299–311.

Summers, Anita A., and Barbara L. Wolfe, "Do Schools Make a Difference?" *American Economic Review*, Vol. 67, No. 3, 1977, pp. 639–652.

Wayne, Andrew J., and Peter Youngs, "Teacher Characteristics and Student Achievement Gains: A Review," paper presented at the annual meetings of the Association for Public Policy and Management, Washington, D.C., 2001.

About the Authors

CHRISTOPHER JEPSEN

Christopher Jepsen is a research fellow at the Public Policy Institute of California, where he studies student demographics, student achievement, teacher labor markets, and school choice. He also has research interests in demographic economics and law and economics. He received a B.A. in mathematics from Grinnell College and a Ph.D. in economics from Northwestern University.

STEVEN RIVKIN

Steven Rivkin is an associate professor of economics at Amherst College. He is a specialist in labor economics whose areas of expertise include the economics and sociology of K–12 education and race and gender differences in schooling, earnings, and employment. In recent work, he has investigated changes in education productivity, differences in school and teacher quality, teacher labor markets, school desegregation, peer group influences, the efficacy of special education programs, student mobility, and charter schools. He holds a B.A. in economics from the University of Michigan and an M.A. and Ph.D. in economics from the University of California, Los Angeles.

Other PPIC Publications on Education

The Changing Role of Education in the California Labor Market
Julian R. Betts

Equal Resources, Equal Outcomes? The Distribution of School Resources and Student Achievement in California
Julian R. Betts, Kim S. Rueben, Anne Danenberg

"At Home and in School: Racial and Ethnic Gaps in Educational Preparedness"
California Counts: Population Trends and Profiles
Volume 3, Number 2, November 2001
Jennifer Y. Cheng

Math Matters: The Links Between High School Curriculum, College Graduation, and Earnings
Heather Rose, Julian R. Betts

For Better or For Worse? School Finance Reform in California
Jon Sonstelie, Eric Brunner, Kenneth Ardon

School Finance and California's Master Plan for Education
Jon Sonstelie, Peter Richardson (editors) et al.

"The Linguistic Landscape of California Schools"
California Counts: Population Trends and Profiles
Volume 3, Number 4, February 2002
Sonya M. Tafoya

PPIC publications may be ordered by phone or from our website
(800) 232-5343 [mainland U.S.]
(415) 291-4400 [Canada, Hawaii, overseas]
www.ppic.org